SEMI-SYNTHETIC PROTEINS FOR CATALYTIC AND ANALYTICAL APPLICATIONS

A Thesis
Presented to
The Academic Faculty

by

Karl J. Hüttinger

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Chemistry

Georgia Institute of Technology

May 2009

SEMI-SYNTHETIC PROTEINS FOR CATALYTIC AND ANALYTICAL APPLICATIONS

Approved by

Dr. Christoph J. Fahrni, Advisor

School of Chemistry and Biochemistry

Georgia Institute of Technology

Dr. Andreas S. Bommarius

School of Chemical and Biomolecular Engineering

Georgia Institute of Technology

Dr. Angus P. Wilkinson

School of Chemistry and Biochemistry

Georgia Institute of Technology

Dr. Z. John Zhang

School of Chemistry and Biochemistry

Georgia Institute of Technology

Dr. James A. Mulholland

School of Civil & Environmental Engineering

Georgia Institute of Technology

Date Approved: January 23, 2009

ACKNOWLEDGMENTS

I would like to thank Dr. Christoph J. Fahrni for his guidance during my graduate studies at Georgia Tech. I would also like to thank Reagan McRae for her help on biological experiments but especially for her moral support and friendship. Moreover, I would like to thank Dr. Subrata Mandal, Dr. Yonggang Wu, Dr. John Cody and all others who worked in this group. Sybille Krause helped me to get into Georgia Tech and the importance of her support cannot be underestimated. I am very grateful for the constant support and trust of my family despite being separated by the Atlantic Ocean. Through my experiences in graduate school, I have finally learned to understand the true meaning of the following Latin saying that I memorized a long time ago:

“Per aspera ad astra!”

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Å	Angstrom
AAS	atomic absorption spectroscopy
AGT	O^6-alkylguanine-DNA alkyltransferase
b	broad signal
C	Celsius
Cer	Cerulean
cm	centimeter
CD	circular dichroism
CT	charge transfer
D	doublet
deg	Degree
DNA	Deoxyribonucleic acid
E.coli	Escherichia coli
EI	electron ionization
ESI	electrospray ionization
FID	Flame Ionization Detector

FP	Fluorescent Protein
FRET	Förester Resonance Energy Transfer
HOMO	highest occupied molecular orbital
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
Hz	Hertz
ICPMS	inductively coupled mass spectrometry
IR	infrared
K	Kelvin
Λ	Wavelength
LUMO	lowest unoccupied molecular orbital
m	multiplet
MHz	Megahertz
mm	millimeter
MS	mass spectrometry
nm	nanometer
NMR	nuclear magnetic resonance
ppm	part per million

q	quartet
s	singlet
TGN	*trans*-Golgi network
t	triplet
THF	tetrahydrofuran
UV	ultraviolet

SUMMARY

Proteins have evolved over millions of years to serve a plethora of highly specialized functions in biological systems. Given the enormous diversity in structure and function, it is truly surprising that only 20 different amino acids are utilized as the building blocks of proteins. Furthermore, only a small set of metal cations that are biologically available are used as structural or catalytically active cofactors in proteins, whereas rare metal cations such as platinum, ruthenium or rhodium remain absent. In the 20th century myriad catalysts, based on non-biological transition metals, emerged that can facilitate numerous organic transformations. The goal of my thesis was to introduce new functions into proteins by attaching platinum metals and fluorescent metal sensors. Thus, semi-synthetic proteins for catalytic and analytical applications were generated.

Chapter I provides the background for the generation and development of semi-synthetic proteins for catalytic and analytical applications. This chapter also provides background about challenges in sustainable chemistry, water as a reaction medium for organic reactions and catalysis, the role of copper, iron, and zinc ions in biology, small molecule fluorescent probes for visualization of biologically relevant metals and fluorescent proteins as tools for cellular imaging. Chapter II and III focus on semi-synthetic proteins for catalytic applications while chapter IV focuses on their analytical applications.

The replacement of organic solvents by environmentally benign solvents such as water is an imperative step towards achieving “green chemistry”. The combination of small molecule catalysts with proteins may introduce new functions and take advantage

of the benefits of “both worlds” while avoiding their potential drawbacks. Therefore semi-synthetic catalysts were developed for enantioselective organic reactions in aqueous medium.

Chapter II discusses the design, development and characterization of a suitable reaction, reaction conditions and catalytic system for later utilization in a semi-synthetic protein. Ruthenium porphyrins catalyzed cyclopropanation reactions with fair yields and high stereoselectivity in aqueous medium. The successful reaction in water was a crucial requirement for a catalytically active semi-synthetic protein. Mechanistic studies did not elucidate the actual catalytic species for the formation of the cyclopropanation product and the side-product diethyl maleate; however, new insights were gained from the analysis of potential reaction pathways. Moreover, studies of the influence of axial ligands, resembling likely residues coordinating to the ruthenium metal center in the active site of a semi-synthetic protein, on the carbene formation of ruthenium porphyrins illustrated that coordination of axial ligands may inhibit the catalytic activity.

In chapter III, the generation of ruthenium porphyrin based semi-synthetic proteins and their subsequent catalysis of cyclopropanation reactions is discussed. Myoglobin and myoglobin mutants were successfully reconstituted with a heme-like ruthenium carbonyl porphyrin; however, none of the formed semi-synthetic proteins catalyzed the enantioselective cyclopropanation of styrene. Efforts to determine the reconstitution efficiency of the generated semi-synthetic were hampered by problems to purify the generated semi-synthetic proteins that are probably due to non-specific binding of the ruthenium porphyrin to the protein surface.

The exploration of labile metal pools of the biologically relevant transition metals copper, iron and zinc in cells was the goal of developing semi-synthetic proteins for analytical applications. Combining fluorescent proteins with colored or fluorescent metal chelators by forming semi-synthetic proteins allows taking advantage of their beneficial properties while avoiding their downsides. This design offers an attractive platform for *in vivo* metal sensing.

Chapter IV discusses the design and generation of semi-synthetic proteins for analytical applications. Plasmids encoding fluorescent proteins, targeting sequences and AGT or intein fusion domains (necessary for labeling) for eukaryotic and prokaryotic expression were generated. The targeting of intracellular compartments (mitochondria, nucleus and TGN) was successful (confirmed by light microscopy experiments with transfected mammalian cells). *In vitro* labeling experiments of expressed and purified fusion proteins with rhodamine derivatives succeeded with AGT based fusion proteins; however, labeling of fusion proteins by *trans*-splicing with split-inteins failed. A zinc(II)-chelator was attached to an AGT based protein and the resulting semi-synthetic protein exhibited strong changes of fluorescence in the presence of zinc(II). This represents an important step towards the goal of *in vivo* cell imaging of labile zinc(II) pools. Iron chelators suitable for protein-labeling could not be synthesized. Despite extensive efforts, all attempts failed to generate a chelator that forms Cu(I)-complexes with the 1:1 stochiometry (ligand:metal) that is necessary for metal sensing with semi-synthetic proteins.

Chapter I

INTRODUCTION

1.1 Semi-Synthetic Proteins for Catalytic Applications

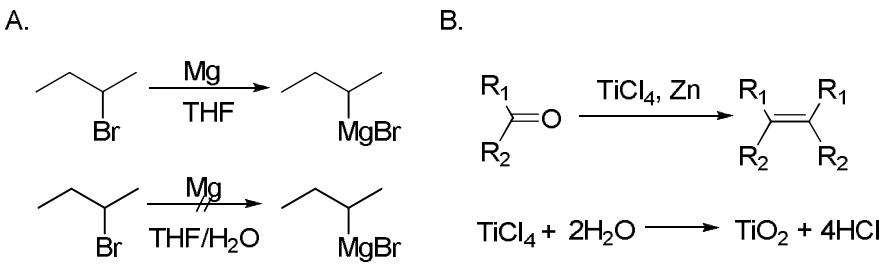
1.1.1 Green Chemistry and Its Importance

In 1990 the Pollution Prevention Act was passed in the United States, representing the beginning of Green Chemistry.[1,2] Since then the Environmental Protection Agency (EPA) has promoted Green Chemistry by grants, education, partnerships, research, and the notable Presidential Green Chemistry Challenge Awards.[3] Green Chemistry, also called sustainable chemistry, is based on a set of principles that reduces or eliminates the use or generation of hazardous substances in the design, manufacture and application of chemical products.[4] The expenditure on environmental protection has risen in western industrialized countries because of increasingly stringent environmental protection regulations.[5] Therefore it is possible to get a competitive advantage by applying green chemistry technologies.[6] The significance of green chemistry was recognized by awarding Richard R. Schrock, Robert H. Grubbs, and Yves Chauvin with the Nobel Prize in Chemistry in 2005 for the development of the metathesis method in organic synthesis. The Nobel Prize Committee stated "this represents a great step forward for 'green chemistry', reducing potentially hazardous waste through smarter production. Metathesis is an example of how important basic science has been applied for the benefit of man, society and the environment."[7]

There are several principles of green chemistry, stating for example that “catalytic reagents are superior to stoichiometric reagents.” or that auxiliary substances like solvents should be “innocuous when used”.[4] Catalysis in aqueous medium fulfills these criteria and therefore represents an attractive way to develop methods for sustainable chemistry.

1.1.2 Reactions in Aqueous Media

Water is a highly polar solvent with a high dielectric constant (ε_r = 78.30) compared to common organic solvents like ethanol (ε_r = 24.55), diethyl ether (ε_r = 4.20), or n-hexane (ε_r = 1.88). It is unique among solvents because of its exceptionally small molecular volume and its ability to form a three dimensional hydrogen bonding network with a dominant tetrahedral coordination.[8] Due to these properties most organic compounds show only limited solubility in water. Moreover, water can deactivate many reactions. A classic text book example is the deactivation of the reactive magnesium surface for the formation of the Grignard reagent (Scheme 1-1).[9] Besides, water has a deleterious effect on many reagents and organic compounds. In the McMurry reaction even trace amounts of water react readily with $TiCl_4$ in an exothermic fashion.[10,11]



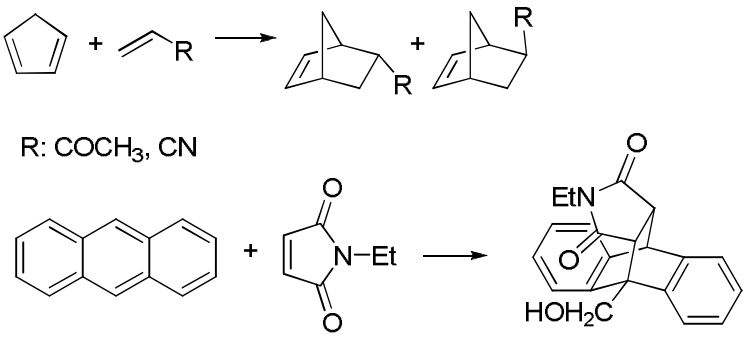
Scheme 1-1.[10,11]

Therefore it is no surprise that water as a solvent was ignored for a long time or even considered a contaminant in organic reactions.[12]

Nevertheless, water has several properties that make it attractive for synthetic applications. First, water is the cheapest solvent and readily available. Consequently, the utilization of water can make many chemical processes more economic. Besides, the large heat capacity of water, which is one of the largest ones of all substances, allows easier control of the reaction temperature.[13] Moreover, in certain large scale industrial processes isolation of the organic products can be achieved by simple phase separation.[14] The innocuous water has crucial advantages with regards to safety, because many organic solvents constitute health hazards by being inflammable, potentially explosive, mutagenic, and/or carcinogenic.[15] The chemical industry is a major producer of environmental pollution. The development of nonhazardous solvent alternatives like water is of great importance in order to respond to increasing regulatory pressure focusing on organic solvents.[16] Water's high polarity and ability to form a hydrogen-bonding network can lead to a significant increase of the reaction rate.[17-19] These properties can also result in beneficial effects on organic reactions regarding selectivity.[20] Lengthy protection-deprotection steps can also be reduced in case of the presence of acidic moieties or by using water-soluble compounds without derivatization with hydrophobic moieties.[12] Thus, the possible benefits of water as a reaction medium has prompted the reevaluation of reactions that were optimized for common solvents like dichloromethane, dimethylsulfoxide or hexanes.

The discoveries made in the early 1980s by the research groups of Breslow[21-23] and Grieco[24-26] on the dramatic effect of water on the reaction rates and

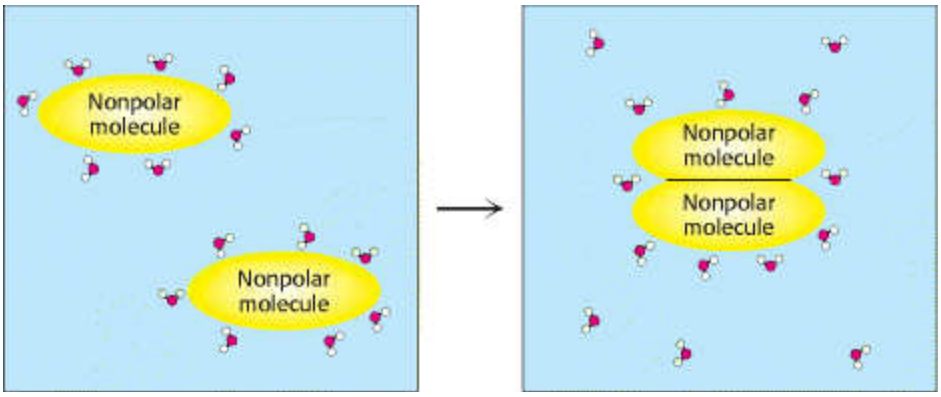
stereoselectivities of the Diels-Alder reaction (Scheme 1-2) triggered a broad interest in the field of aqueous organic synthesis.



Scheme 1-2.[23]

The strong enhancement of the reaction rate in water, compared to solvents like acetonitrile, is a phenomenon that cannot be simply explained by the higher polarity of the aqueous media. Addition of lithium chloride, a “salting-out” agent, decreases the solubility of hydrocarbons in water. The lowered solubility resulted in an even higher rate enhancement compared to plain water. On the other hand salts like guanidinium chloride, which are known to solubilize hydrocarbons in water, decreased the reaction rate. These findings can only be explained by the hydrophobic effect (Scheme 1-3). The hydrophobic effect is based on hydrophobic interactions[27,28] between nonpolar moieties molecules. This process, in which water forces nonpolar molecules to associate, is crucial for biological systems[29,30] (e.g. protein folding) but was mostly ignored for synthetic applications. The driving force of the aggregation of hydrophobic molecules is

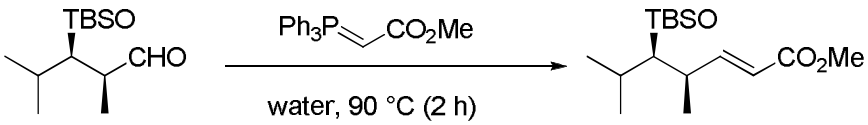
a combination of van der Waals interactions and the increase of entropy by the release of water molecules.



Scheme 1-3. [31]

The great impact of hydrophobic effects on simple organic model reactions like the benzoin condensation in water has been studied in detail since the discovery of rate enhanced Diels-Alder reactions.[32,33]

Poor solubilities of substrates in aqueous media do not necessarily present obstacles for organic synthesis if the hydrophobic effects are strong enough to enable the reaction. The Wittig reaction was carried out in water utilizing poorly water-soluble stabilized ylides to produce enoates that are instrumental templates for the synthesis of macrolide antibiotics (Scheme 1-4).[17]



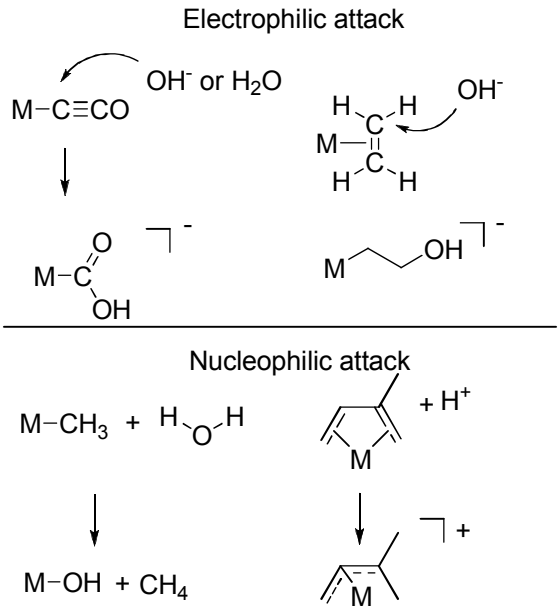
Scheme 1-4.[17]

Despite the reactants and the product being poorly water-soluble the reaction proceeded with nearly the same yield as in CH_2Cl_2. However, the reaction time was only two hours in water, in contrast to four weeks in CH_2Cl_2.

Sharpless and coworkers demonstrated that water is an ideal medium for many organic transformations, like “click chemistry” or the preparation of substituted tetrazoles, despite poorly soluble starting materials and products.[34,35] Hence, poor solubility in water is not simply an obstacle but, contradictory to previous assumptions in the scientific literature, can even be beneficial.

1.1.3 Organometallic Reactions in Water by Small Molecule Catalysts

A plethora of chemical reactions can be facilitated utilizing small molecule organometallic catalysts;[36-40] however, the vast majority of reactions is conducted in organic solvents. Apart from the general difficulties of aqueous organic chemistry (Chapter 1.2), the challenges of organometallic reagents lie in the thermodynamic instability of the metal-carbon bond with regard to their hydrolysis products (Scheme 1-5).[8,41]



Scheme 1-5.[41]

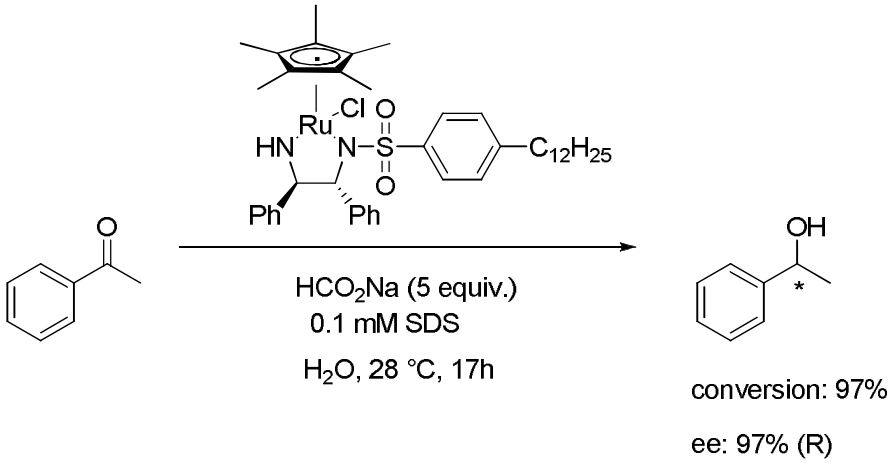
Moreover, water can act as a ligand itself and replace ligands that are bound to the metal center. Hydrolysis reactions do not always deactivate or destroy the catalyst. In some reactions (e.g. Wacker-Hoechst acetaldehyde process) water may even form the catalytically active species.[42]

In spite of these difficulties many organometallic reactions in aqueous media have already been developed.[43-53] Faster reaction to the desired products than decomposition or deactivation by water allows these reactions to be carried out in aqueous medium (kinetic control). The most important industrial reaction utilizing aqueous organometallic catalysis is the hydroformylation reaction of alkenes.[54] Rhodium complexes are used for the production of more than 400,000 tons of n-Butanal per year.[55]

1.1.4 Enantioselective Organometallic Reactions in Water by Small Molecule Catalysts

The development of methods to produce enantiomerically pure molecules is of critical importance, especially in the pharmaceutical industry. Thalidomide is a racemic drug that was sold in nearly 50 countries from 1957 to 1961 as a sleeping aid and to combat morning sickness of pregnant women. The R(+) enantiomer had a sedative effect while the S(-) enantiomer is teratogenic and caused serious birth defects.[56] Hence, the enantiopurity of drugs has become a crucial property.

Myriad small molecule organometallic catalysts have been developed for enantioselective transformations of compounds in organic solvents;[36-40] however, with a few notable exceptions, the development of chiral ligands and asymmetric catalysts was conducted almost exclusively in organic solvents. By the recent focus on aqueous organometallic catalysis several reactions with high stereoselectivity and yield have been developed.[12,46,51,53,57] A notable example is the rhodium catalyzed asymmetric transfer reaction of alkyl and aryl ketones in aqueous media by Adolfsson and coworkers (Scheme 1-6).[58]



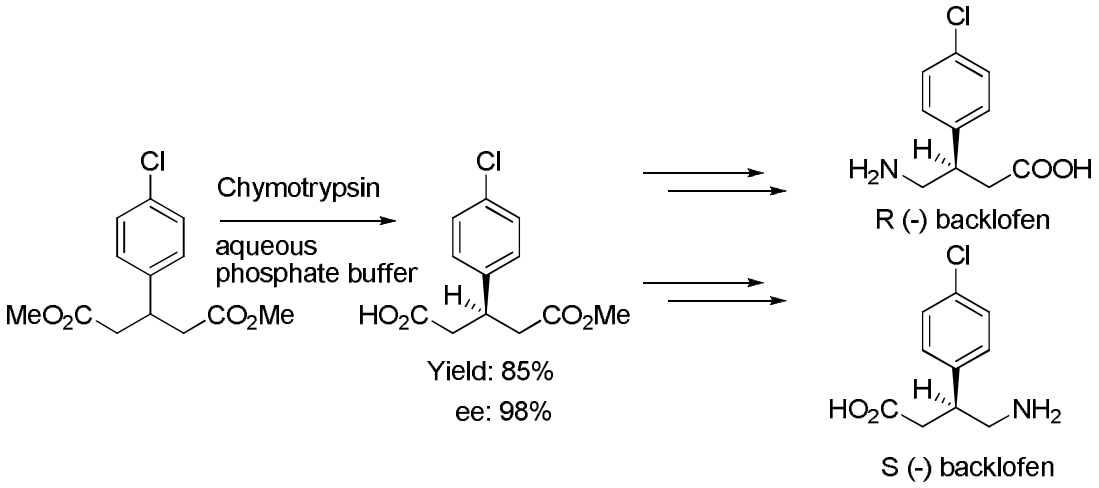
Scheme 1-6.[58]

The rhodium catalyst is a modified catalyst that was originally developed by Noyori and coworkers and is formed in situ.[59] The reduction to the corresponding alcohol was facilitated with sodium formate as hydride donor and SDS as surfactant. Several alkyl and aryl ketones were reduced with an enantiomeric excess of up to 97% at a conversion of up to 99%.

1.1.5 Enzymes and Enzyme Mimics

The synthesis of enantiomerically pure compounds necessitates chiral organometallic catalysts, which require often multi-step preparations and are therefore too expensive for large scale industrial applications.[60,61] An alternative to small molecule organometallic catalysts for enantioselective organic transformations is the utilization of enzymes. Enzymes are biological, protein-based catalysts that use water as their natural solvents. Myriad enzymatic reactions and their corresponding enzymes have been characterized.[62] These biocatalysts usually catalyze reactions with very large turnover

numbers and high regio- and stereoselectivity.[31] Enzymes are typically easy to handle due to their stability under ambient conditions, unlike small molecule organometallic catalysts that tend to decompose in air, by moisture, and by light. Moreover, the biodegradability of enzymes can help to reduce the amount of produced toxic waste. Besides, properties like stereospecificity and turnover numbers for a particular substrate can be enhanced by mutations of the active site. Due to their lack of stability in organic solvents and high temperatures enzymes are often only used in industry for reactions that require highly selective catalysts. In particular enzymes are utilized for enantioselective transformations in the synthesis of drugs.[63-69] Baclofen for instance is an antispastic drug, which was sold as a racemate despite significantly different pharmacological and toxicological properties of the enantiomers.[69,70] The chemoenzymatic synthesis of both enantiomers was carried out in order to solve this problem (Scheme 1-7).[71]



Scheme 1-7.[71]

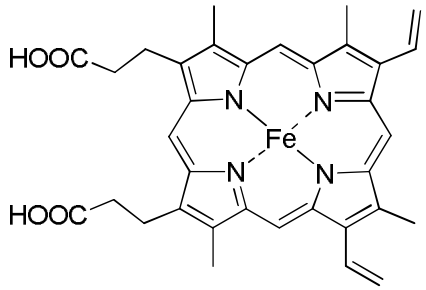
The key step for the synthesis of the enantiomers was the enantioselective hydrolysis of the prochiral dimethyl 3-(p-chlorophenyl)glutarate to its half-ester by the enzyme chymotrypsin.

1.1.6 Semi-Synthetic Proteins for Catalysis

Most enzymes used in organic synthesis offer large turnover number and high stereoselectivity but exhibit a narrow substrate tolerance while small molecule organometallic catalysts usually accept a much broader range of substrates.[72] A viable approach to combine the benefits of "both worlds" lies in the design of semi-synthetic proteins. Semi-synthetic proteins consist of a protein and an incorporated synthetic prosthetic group.

Several research groups have focused on the incorporation of organometallic cofactors into proteins and the characterization of the resulting semi-synthetic proteins. The prosthetic groups of many heme proteins were successfully removed and reconstituted with either the native or modified heme group.[73] Hayashi, Lu, and Watanabe and coworkers have utilized the heme-protein Myoglobin as their model-protein for the introduction of prosthetic groups.[74-113] Myoglobin plays an essential role in the storage of oxygen during periods of rest until required for oxidative phosphorylation. It is a small single-chain globular protein with a molecular weight of about 18 kDa and contains the porphyrin cofactor iron protoporphyrin IX (Figure 1-1 and Figure 1-2). Most commonly used are horse heart myoglobin and sperm whale myoglobin, which amino acid sequence differs only in 18 positions from each other. These differences appear mainly at the opposite side of the heme edge but not in the active site itself.[114]

Figure 1-1. Cartoon rendering of X-ray structure of sperm whale myoglobin including heme group that is depicted as ball and stick structure (PDB entry 1mbo).[115]



heme group = Fe Protoporphyrin IX

Figure 1-2. The native heme group of Myoglobin.

Non-native metal-containing cofactors can be integrated into proteins by non-covalent, single-point and dual-point attachment. Non-covalent attachment is based on hydrogen-bonding interactions, hydrophobic interactions and coordinative bonds to the metal center of the cofactor. The non-native prosthetic groups incorporated by this method include metal complexes for electron-transfer studies, artificial receptors for modulation of heme protein activities and binding domains for protein-protein and protein-small molecule recognition.[76,116] Hayashi and coworkers utilized non-convalent attachment in order to incorporate heme-groups containing artificial interfaces into myoglobin (Figure 1-3 and Figure 1-4).[76] The resulting semi-synthetic proteins show a much higher activity in the sulfoxidation reaction of 2-methoxyphenol than native myoglobin.[107]

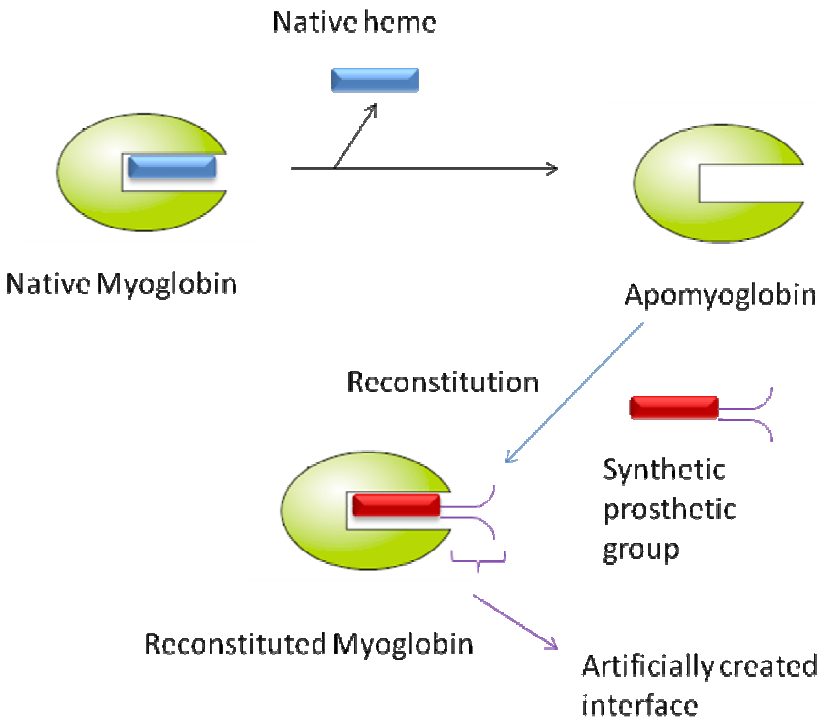


Figure 1-3. Reconstitution of Myoglobin with synthetic prosthetic cofactor based on previous removal of native heme. [76]

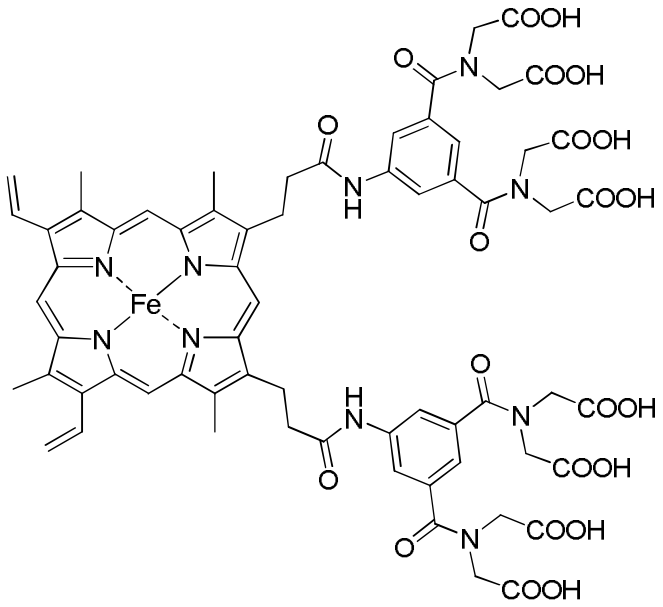


Figure 1-4. Synthetic prosthetic group for generation of an interface. [76]

Both the Lu and the Watanabe research group have focused on the incorporation of metal-salen and salophen derivatives into myoglobin (Figure 1-5).

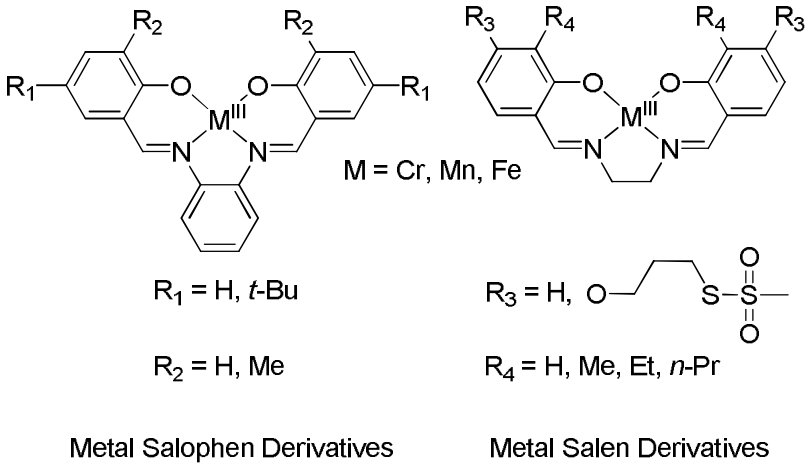


Figure 1-5. Salen and Salophen derivatives utilized for sulfoxidation reactions.[84,100,109]

However, they applied different designs and methods to incorporate these types of organometallic complexes in myoglobin. The Watanabe group utilized non-covalent attachment (Figure 1-6) while Lu and coworkers preferred single-point and dual-point attachment (Figure 1-7).[84,109] The sulfoxidation of thioanisole yielded only low enantiomeric excess (up to 32%) with Watanabe's salen and salophen based modified myoglobin (optimized by H64D/A71G mutation). On the other hand, Lu and coworker's approach to covalently attach salen derivatives to myoglobin (M72C/Y103C) in two positions by disulfur bonds yielded a more stable and rigid attachment and resulted in higher stereoselectivity (up to 51% ee).

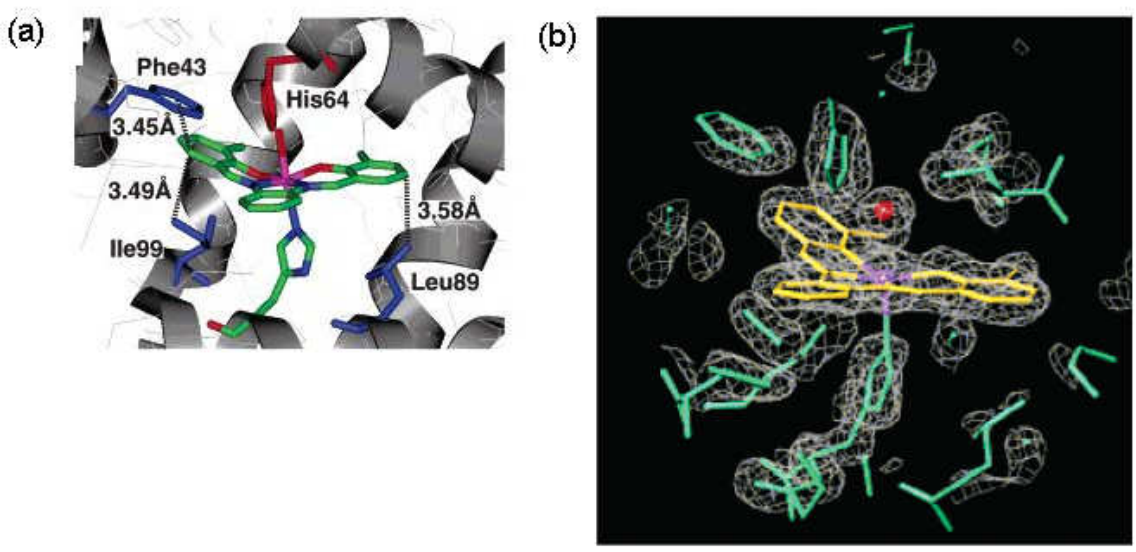


Figure 1-6. (a) Active site structure of Mn(III)-3,3'-Dimethylsalophen A71G apo-myoglobin. (b) Crystal structure of Mn(III)-3,3'-Dimethylsalophen A71G apo-myoglobin.[109]

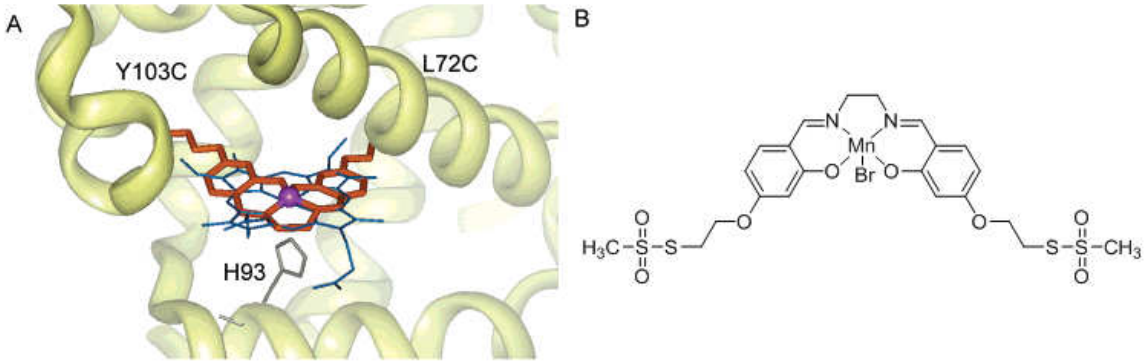


Figure 1-7. (a) Computer model of Mn-salen complex covalently bound to myoglobin backbone. (b) Unbound Mn-salen complex.[84]

Whitesides pioneered the concept of anchoring a homogenous catalyst within a protein environment by non-covalent interactions in order to induce enantioselectivity. A biotinylated diphosphinerhodium(I) complex was incorporated into avidin and enabled the asymmetric hydrogenation of α-acetamidoacrylic acid with up to 44% ee (Figure 1-8).[117] Efficient attachment of the organometallic species is facilitated by the extreme binding affinity of the protein streptavidin to biotin (K_d around 10^{-15}), which is the strongest known non-covalent interaction.[118] This type of supramolecular anchoring approach was optimized by Ward and coworkers. Combinatorial modification of the rhodium organometallic catalyst, the spacer to the biotin moiety and site-directed mutations of streptavidin led to the successful asymmetric hydrogenation of acetamidoacrylic acid with up to 96% ee in favor of the (R) configuration.[119] Remarkably, by applying the same kind of combinatorial modification strategy the enantioselectivity was reversed to 88% ee in favor of the (S) configuration.[120] Recently, new types of reactions (i.e. allylic alkylation and sulfoxidation) were successfully catalyzed by the streptavidin-biotin catalyst attachment approach.[121,122]

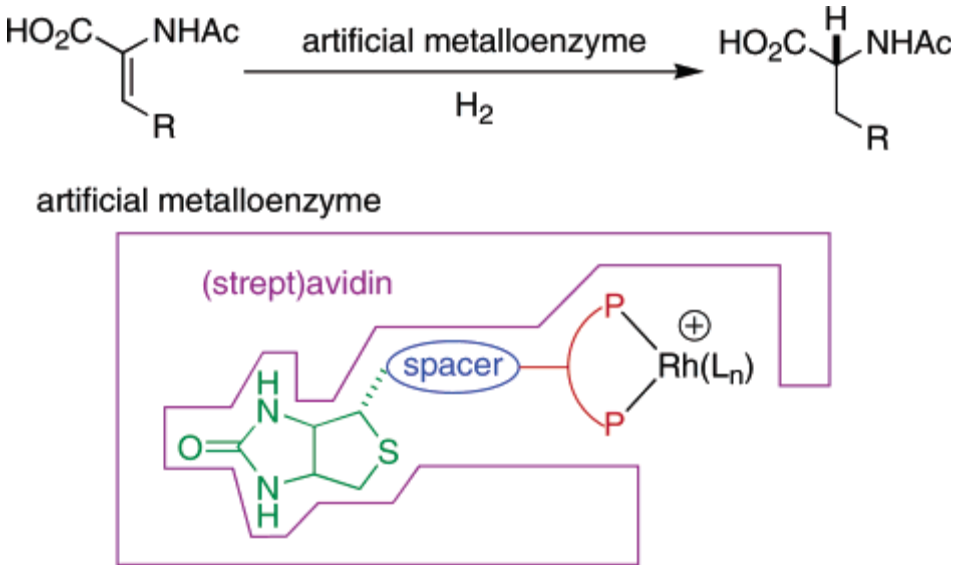


Figure 1-8. Whitesides and Ward design of supramolecular anchoring with (strept)avidin (purple), biotin (green), spacer (blue) and a Rhodium(I) hydrogenation catalyst (red and black). [117,118]

The transport proteins serum albumin, in particular bovine serum albumin (BSA), can be useful for supramolecular anchoring strategies of metal cofactors within a host protein. These exhibit the ability to strongly bind a wide range of hydrophobic guests, including fatty acids, steroids, thyroxine, porphyrins, etc.[118] Several enantioselective transformations were performed with these proteins, including sulfoxidation, epoxidation, reduction, and Diels-Alder reactions, without the presence of an organometallic catalyst.[123-128] Gross and coworkers successfully incorporated amphiphilic bis-sulfonated iron and manganese corroles into various serum albumins through supramolecular anchoring. The in situ formed semi-synthetic protein facilitated the enantioselective sulfoxidation of thioanisole derivatives with up to 74% ee (Figure 1-16).[129] The Mn(III)-corrole conjugate shows a 1:1 conjugate with a K_d in the high nanomolar range. The presence of an induced circular dichroism signal suggests that the metal complex is

integrated into a well-defined chiral environment.[130] This finding is consistent with the X-ray structure of a hemin human serum albumin conjugate. Reetz and coworkers applied the same approach in order to incorporate a water-soluble copper(II)-phtalocyanine into BSA and HSA. The resulting conjugate catalyzed the Diels-Alder reaction of multiple substrates in a highly enantioselective fashion.[131]

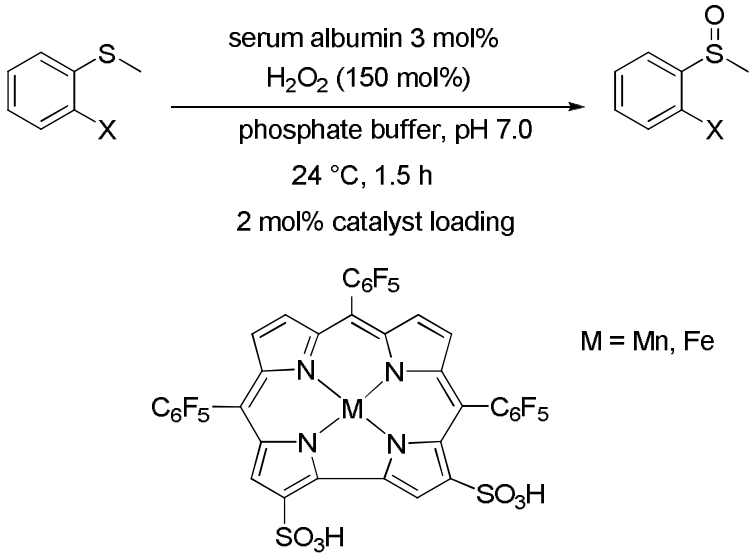


Figure 1-16. Sulfoxidation by semi-synthetic serum albumin conjugates.

Other notable contributions to the field of semi-synthetic proteins include covalent anchoring of Cu-phenanthroline to adipocyte lipid binding protein,[132] covalent anchoring of Mn-Schiff bases to papain,[133] or supramolecular anchoring of Cu-pyridoxamine to Ribonuclease S.[134] A comprehensive review has been recently published by Steinreiber and Ward.[118]

1.2 Semi-Synthetic Proteins for Analytical Applications

1.2.1 Copper, Iron and Zinc in Biology

Copper, iron and zinc are essential trace elements in all organisms and are crucial for the function of numerous enzymes.[135] Iron ranks first in intracellular abundance, followed by zinc and copper.[136] Copper is found in various enzymes as catalytic cofactor, most notably cytochrome c oxidase and superoxide dismutase. It is involved in many cellular processes such as iron uptake, mitochondrial respiration and a range of enzymatic redox processes.[137] Iron is found in myriad enzymes, often incorporated into heme groups. Most notably are cytochrome proteins (redox reactions) and oxygen binding proteins such as hemoglobin and myoglobin. Non-heme iron containing proteins include ribonucleotide reductase (DNA synthesis) and purple acid phosphatase (hydrolysis of phosphate esters).[138] Zinc is essential for growth and development,[139,140] prostate gland function,[141] reproduction,[142,143] olfactory sensing,[144] the immune system,[145-147] and can be found in numerous enzymes.

1.2.2 Homeostasis of Copper, Iron an Zinc

The free intracellular concentrations of copper, iron, and zinc are low compared to the total cellular level presumably because these free metals can generate deleterious reactive oxygen species (ROS). The transition metals copper, iron and zinc are distributed to numerous proteins while maintaining low free metal concentrations, indicating the presence of a sophisticated uptake, transport and recycling system.[138,148,149] Defects of these complex regulatory systems can lead to serious clinical

conditions, e.g. immune system dysfunction (zinc deficiency),[147] Menkes' disease (copper deficiency),[150] and hemochromatosis (iron overload).[151] The elucidation of copper, zinc and iron homeostasis at a molecular level is crucial for understanding the mechanisms of these diseases.

1.2.3 Labile Pools of Copper, Iron and Zinc

High binding affinities of native metal binding proteins are the cause of the low intracellular concentrations of free copper, iron and zinc. It has been shown by uptake studies with radioactive ^{64}Cu(II) that copper uptake occurred rapidly and reached a steady state concentration after about 40-60 minutes.[152] However, upon reaching the steady state level, fast efflux of copper was detected after incubation with copper deficient media within the first 15 minutes. These data contradict the high binding affinities of the native cellular metal binding proteins. Therefore, based on similar results for iron and zinc, a kinetically labile pool was suggested that allows rapid uptake and release of these metal ions.[153,154] This accessible labile pool could be due to a metal exchange between two ligands according to an associative mechanism, which has been shown for the dynamic exchange of zinc.[155] Another possibility is that various weakly binding ligands (in comparison to the native binding proteins) in the cytoplasm such as organic anions (phosphates and carboxylates), surface components of membranes (phospholipid head groups), etc. could act as metal buffers.[156] Storage of copper, iron and zinc in intracellular compartments would also prevent deleterious effects by generation of ROS.[157,158] In case of higher concentration of these metals inside of the compartments than in the cytosol an active transport system against the concentration

gradient must be present. Despite the wealth of information available about labile metal pools in cells many questions remain unanswered especially about the exact nature and subcellular location of labile pools of copper, iron, and zinc.

1.2.4 Chemical Sensing of Metal Ions with Small Molecule Fluorescent Probes

Metal ions can be detected by small molecule metal chelators that emit fluorescence or change the fluorescent properties upon binding. There are usually two types of fluorescent probes for metal ions, “turn on” or ratiometric. “Turn on”-type sensors such as Calcium Green (Ca(II)-sensor) exhibit strong fluorescence or fluorescence enhancement upon binding without a shift (Figure 1-17). Probes such as Fura-2 (Ca(II)-sensor) display spectral shifts upon binding.[159] These kinds of probes are called wavelength-ratiometric because the analyte concentration can be determined by the ratio of fluorescence intensities at different excitation or emission wavelengths.[160] There are many more fluorescent zinc sensors than iron and copper sensors because the presence of copper and iron tends to quench the fluorescence emission of most fluorophores. Of the few examples of copper selective fluorescent probes most are selective for Cu(II). Unfortunately, copper will be present as Cu(I) in the reductive intracellular environment, thus rendering Cu(II) probes unsuitable for cell imaging applications. Various fluorescent small molecule metal probes for zinc and copper are depicted in Figure 1-18. Iron is present in the cell in its +2 and +3 oxidation state. Due to the fluorescence quenching iron is usually measured indirectly by linking an iron selective chelator with a fluorophore (Figure 1-19).[156] The analyte concentration can be determined from the fluorescence quenching of the fluorophore-chelator conjugate such

as FL-Phen (fluorescein-phenantroline) and NBD-DFO (nitrobezodiazole-deferrioxamine) upon binding of iron (Figure 1-19).[161,162]

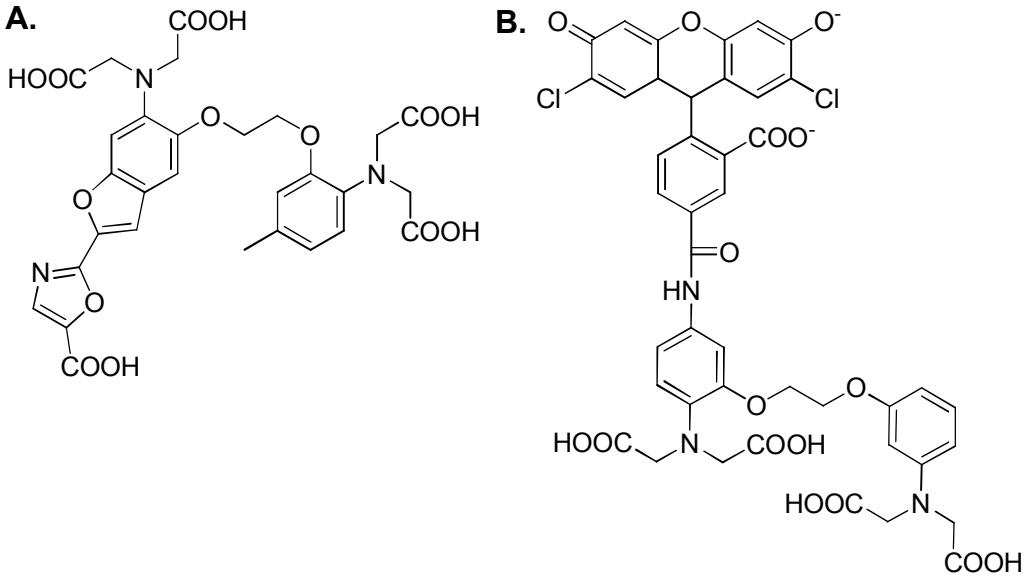


Figure 1-17. Fluorescent calcium probes. **A.** Fura-2 (ratiometric probe). **B.** Calcium Green ("turn-on" probe). [159]

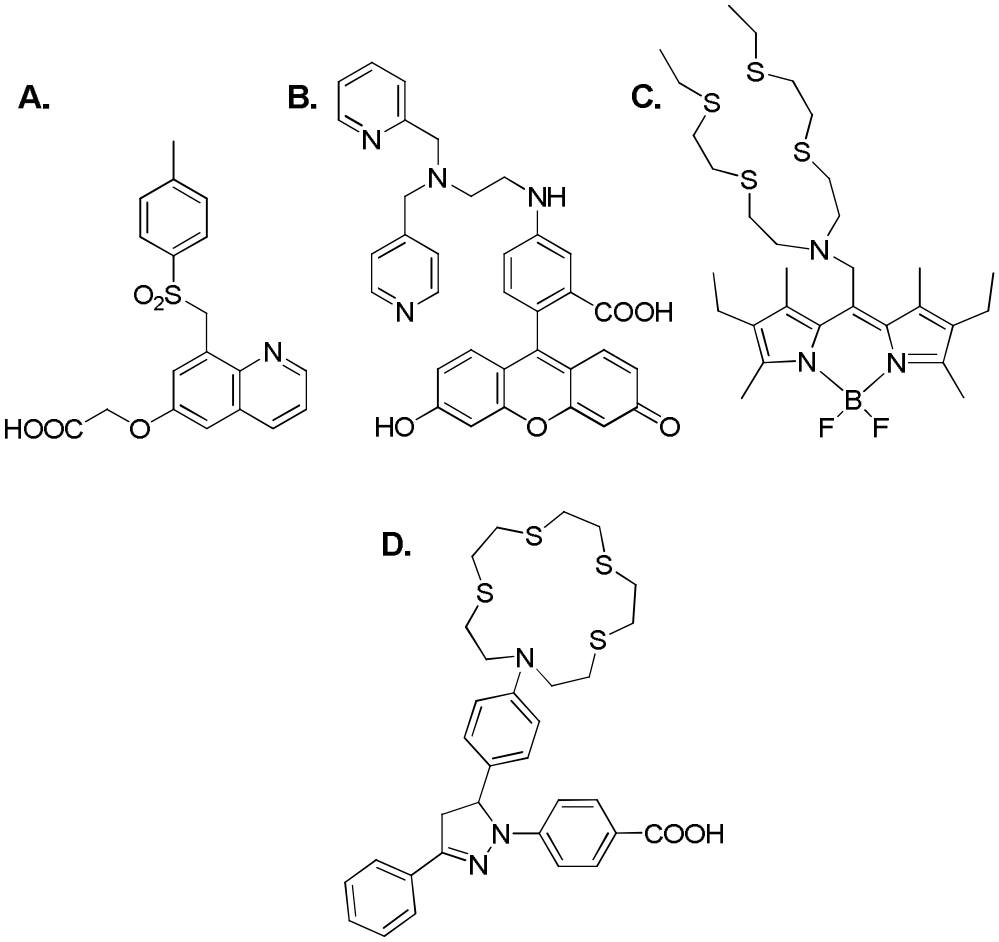


Figure 1-18. Chemical structures of fluorescent copper(I) and zinc(II) probes. **A.** Zinquin (ratiometric Zn(II)-probe).[163] **B.** ZnAF-1(ratiometric Zn(II)-probe).[164] **C.** CS1 ("turn-on" Cu(I)-probe).[165] **D.** CTAP-1 (ratiometric Cu(I)-probe).[166]

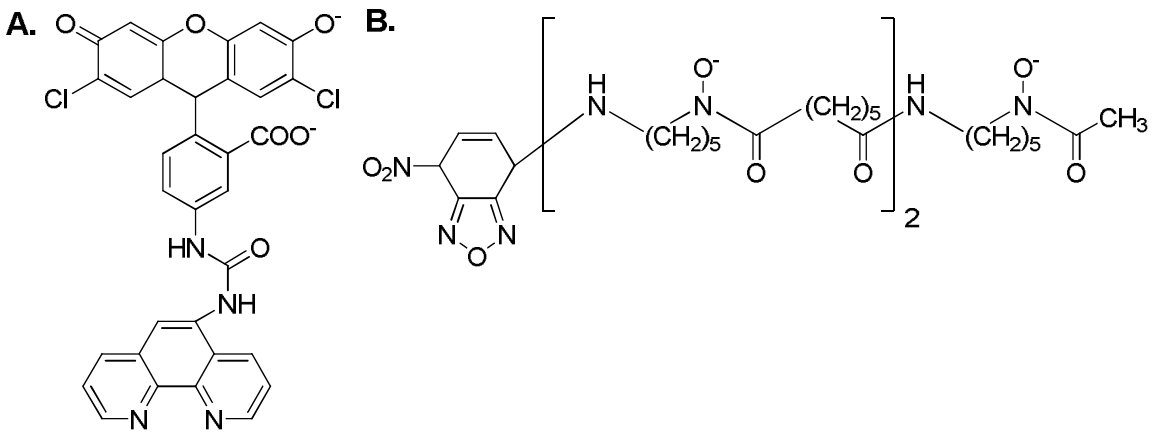


Figure 1-19. Chemical structures of iron probes based on fluorescence quenching of the fluorophore moiety upon binding of iron to the metal chelating moiety **A.** Fl-Phen. **B.** NBD-DFO. [161,162]

1.2.5 Green Fluorescent Protein – A Vital Tool for Cell Imaging Applications

Fluorescence has long been used for cell imaging applications.[167] Originally, fluorescence imaging of biological systems was carried out with small molecule fluorescent dyes attached to the protein of interest through antibodies. This usually necessitates cell fixation and permeabilization. Moreover, fluorescent dyes were also developed that can recognize certain organelles and molecules in the cell. Green Fluorescent Protein (GFP) offers a non-invasive alternative for imaging of living cells and organisms (Figure 1-20). Many mutants, fluorescing in different colors, have been created (Figure 1-21).[168] The genetically encoded fluorescent proteins, which can be expressed as fusion proteins with the protein of interest, have revealed the location and functions of numerous proteins. The significance of fluorescent proteins was recognized by awarding Martin Chalfie, Osamu Shimomura and Roger Y. Tsien the 2008 Nobel Prize in Chemistry for their discovery and development of the green fluorescent protein.[169]

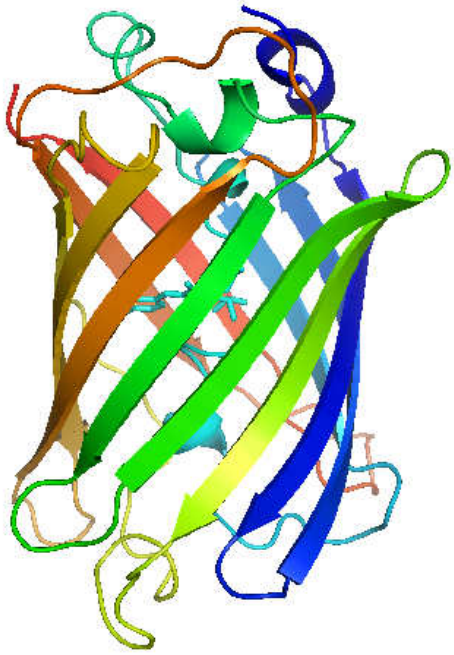


Figure 1-20. Cartoon rendering of X-ray structure of green fluorescent protein (PDB entry 1EMA)

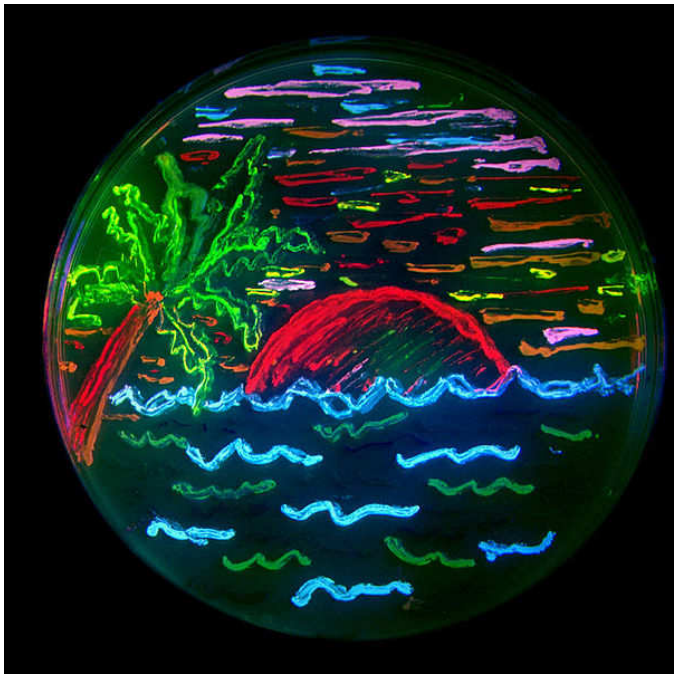


Figure 1-21. San Diego beach scene drawn with *E. coli* expressing 8 differently colored fluorescent proteins (created in Tsien lab).

1.2.6 Signal Peptides

Genetically encoded proteins (e.g. GFP) can be directed to subcellular locations, such as organelles and compartments, via signal peptides. These type of localization signals are short peptide sequences (around 3 - 60 amino acids) that direct the post-translational transport of a protein. Typical targets are peroxisomes (AKL; C-terminal),[170] the nucleus (DPKKKRKV; N or C-terminal),[171] the endoplasmatic reticulum (KDEL; C-terminal)[172] and mitochondria (MSVLTPLLLRGLTGSARRLPVPRAKIHSL; N-terminal).[173] Localization signals are most often located at the N-terminus.

1.2.7 Cell Imaging by FRET and FLIM

Förster Resonance Energy Transfer (FRET) occurs between a donor molecule in the excited state and an acceptor molecule in the ground state.[160] The donor emits at shorter wavelengths (higher energy) and overlaps with the acceptor spectrum. The acceptor is either another fluorophore or colored species. Energy transfer occurs in a nonradiative fashion as a result of dipole-dipole interactions between donor and acceptor. Consequently, the fluorescence emission of the donor decreases upon energy transfer (Figure 1-22). FRET efficiency decays with the sixth power of distance between donor and acceptor. Thus, high FRET efficiency is obtained with sufficiently large spectral overlap, favorable dipole-dipole orientation and close proximity. The strong distance dependence renders FRET a valuable tool for studying protein-protein interactions and changes of protein conformations. The genetically encoded CFP-YFP (cyan fluorescent protein – yellow fluorescent protein) FRET-pair offers good FRET efficiency (if FRET requirements are met) and is widely used for cell imaging applications of proteins and protein-protein interactions.[174]

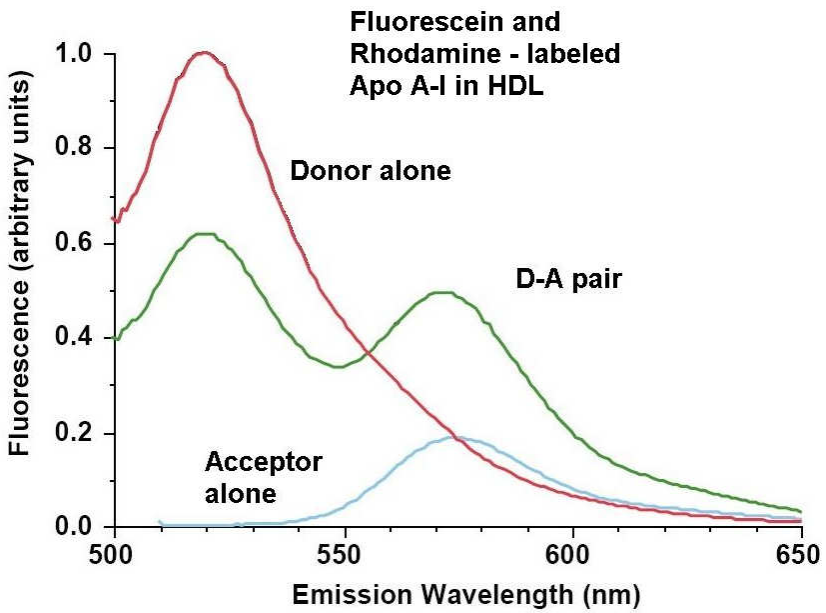


Figure 1-22. Fluorescence emission spectra of FRET pair Fluorescein-Rhodamine in high density lipids.[160]

FRET can also be measured based on the rate of decay of the fluorescence signal rather than fluorescence emission (FLIM = Fluorescence Lifetime Imaging) (Figure 1-23). The advantage of this method is that fluorescence lifetime is unaffected by the probe concentration or excitation intensity.

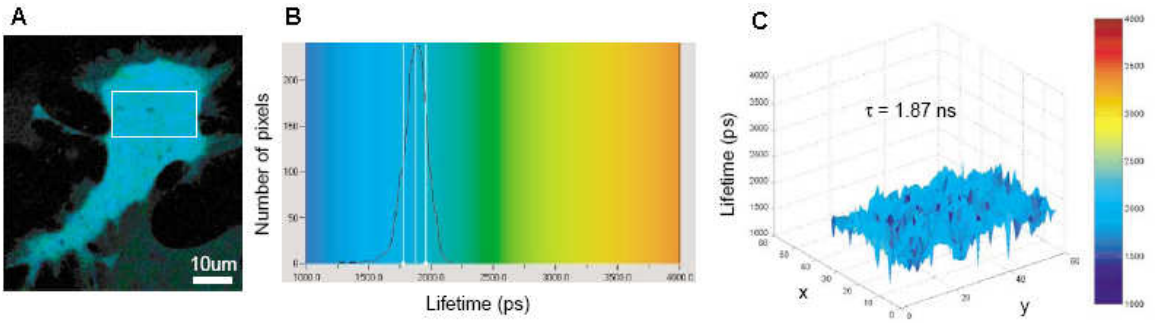
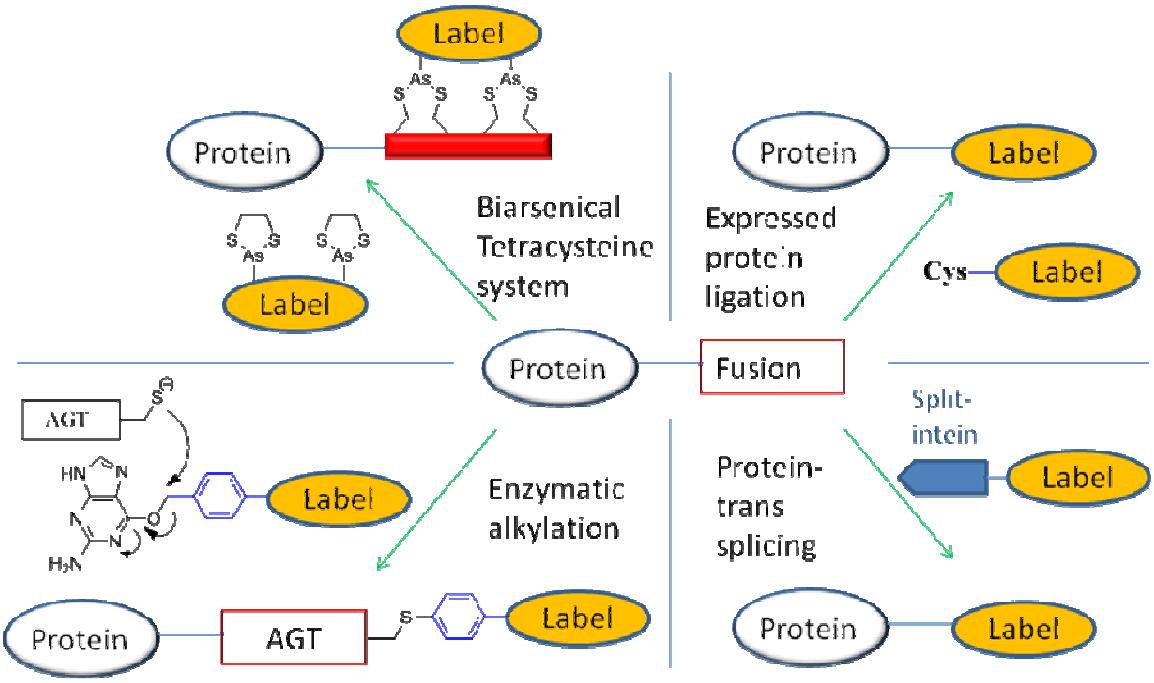


Figure 1-23. Protein localization using FRET-FLIM microscopy of CFP-YFP couple. **A.** Fluorescence lifetime distribution in cell. **B.** Two-dimensional lifetime distribution within specified area. **C.** Three-dimensional plot of lifetime distribution.[175]

1.2.8 Protein Labeling for *In Vivo* Cell Studies

In vitro studies in clearly defined and simplified systems have improved the understanding of many biological aspects and systems. However, the insights gained by *in vitro* studies are limited because an artificial system cannot simulate all physiological characteristics of a cell such as localization, dosage, timing and intermolecular interactions.[176] Utilizing cysteine moieties (naturally occurring or introduced by point-mutations) within the peptide chain of a protein has been a common method to functionalize proteins *in vitro* (usually by formation of a disulfide-bond) but is not suitable for *in vivo* applications due to the abundance of soluble thiols in the intracellular environment. The requirements for successful *in vivo* labeling of proteins are efficient reactivity in aqueous solution, inertness of the label to the surrounding cellular environment, high selectivity or even specificity to the protein of interest and high stability of the formed protein-label adduct.[177] Several methods, which meet the requirements for *in vivo* labeling of proteins, have been developed (Scheme 1-8). All of the illustrated methods are based on the expression of the protein of interest with a

fusion portion that allows reaction with the desired label and will be explained in detail in the following chapters.



Scheme 1-8.

1.2.8.1 Biarsenical Tetracysteine System

In the biarsenical tetracysteine protein labeling approach the label contains two trivalent arsenic ethanedisulfide moieties that bind to a tetracysteine motif of the protein of interest (Cys-Cys-Xaa-Xaa-Cys-Cys; Xaa = non-cysteine amino acid) (Scheme 1-8).[178-180] This method is a particularly useful alternative if the size of GFP (≈ 25 kDa) leads to perturbation of the protein function. Therefore most of the utilized labels are cell-permeable fluorescein and rhodamine derivatives (≈ 0.5 kDa). One of the drawbacks of

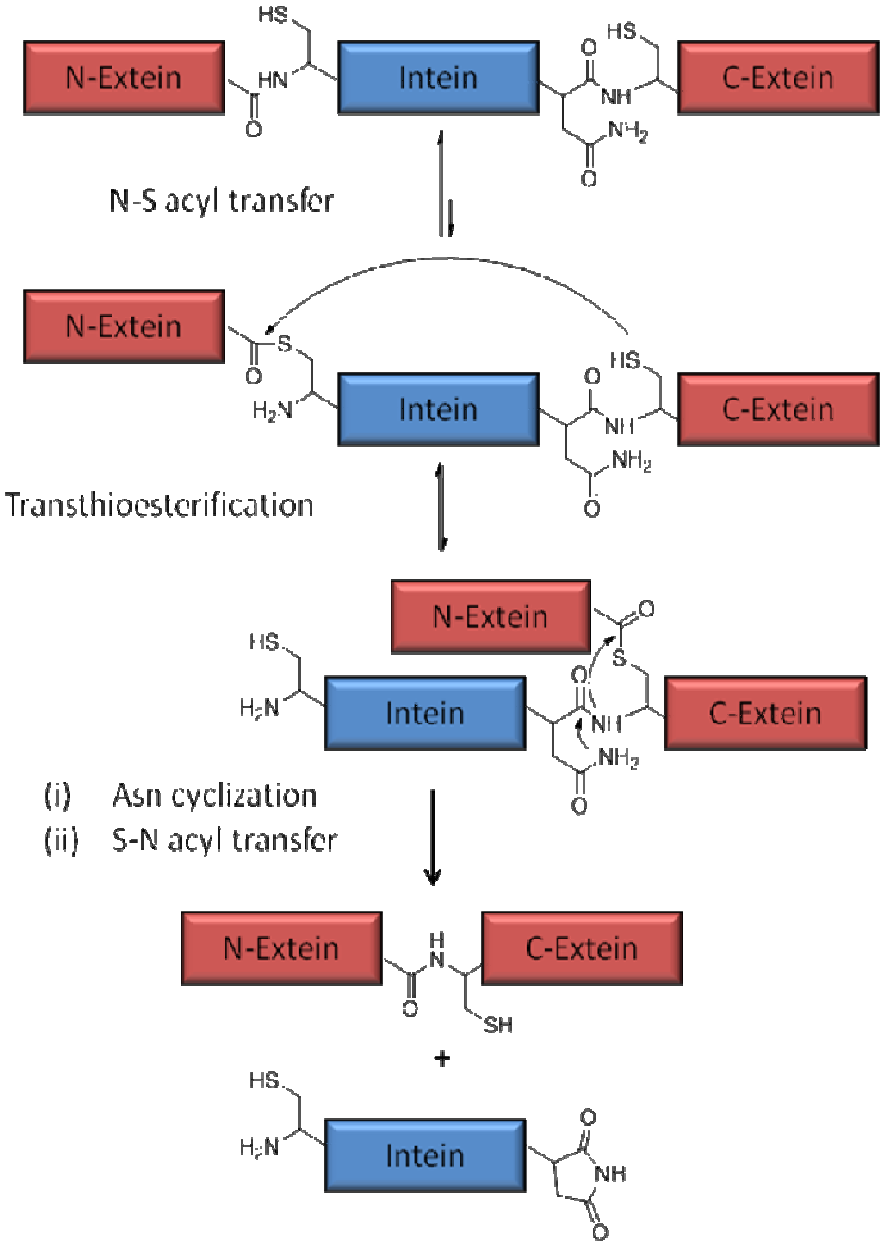
this method is the higher fluorescence detection threshold compared to GFP due to background labeling.

1.2.8.2 Protein Splicing and Trans-*Splicing*

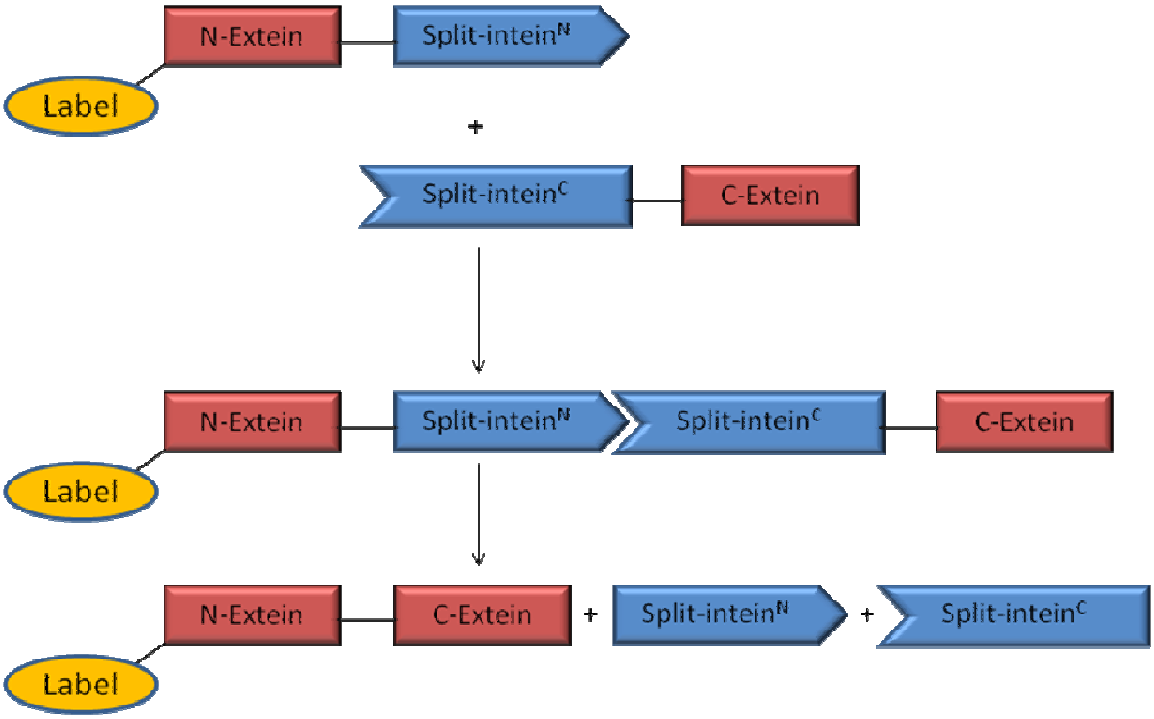
Protein splicing is an autocatalytic posttranslational process in which an intervening polypeptide sequence (intein) is excised while the flanking sequences (exteins) are joined by a peptide bond.[181,182] The intein domain (≈150 amino acids) cannot be excised by itself. Cysteine or serine at the N-terminus and the triad His-(Asn/Gln)-(Ser/Thr/Cys) at the N-terminus of the splicing junction are required for the splicing reaction. The splicing mechanism consists of four major steps (Scheme 1-9). First there is an N-S shift (or N-O shift in case of serine), which is a rearrangement that leads to a thioester (or ester in case of serine). This step is followed by nucleophilic attack of the thiol moiety of the intein flanking cysteine (transesterification; OH moiety in case of serine and threonine). Cyclization of the Asn/Gln residue results in the release of the intein domain. The reaction concludes with an S(O)-N acyl shift of Ser/Thr/Cys to form a peptide bond between the exteins.

Trans-splicing with split-inteins takes advantage of the protein splicing reaction (Scheme 1-10). The intein domain can be cut into two pieces. The splicing reaction is initiated by recombination of the split-inteins. This method can be utilized to label proteins.[183] In theory the label can be attached to the N-extein or C-extein. Connecting the label with the N-extein domain is more feasible in case of generating the label-N-extein-split-intein^{N} portion by synthetic means because the intein^{N} domain can be as little

as 11 amino acids.[184] Long reaction times and reported yields of only about 30% are challenging downsides of this labeling method.



Scheme 1-9.



Scheme 1-10. Mechanism of *trans*-splicing by split-inteins.

1.2.8.3 Expressed Protein Ligation

Expressed Protein Ligation (EPL) is based on the native chemical ligation method, whereby two unprotected synthetic peptides are joined, forming a peptide bond.[185] In native chemical ligation an α-thioester reacts with an N-terminal cysteine specifically even in the presence of internal cysteines. In case of *in vivo* EPL the protein of interest is fused to a modified intein domain (Scheme 1-8). Incubation with a cell-permeable N-cysteine containing label leads to reaction with the expressed protein and displacement of the intein domain.[186] Since N-terminal cysteines do not occur in biological systems naturally side-reactions by competition with N-cysteine containing endogenous proteins can be ruled out.

1.2.8.4 Human Alkylguanine Transferase

Mutant Human O^6-alkylguanine-DNA alkyltransferase (AGT) is a monomeric protein (≈ 177 amino acids) that can be labeled with substrates containing an alkylguanine moiety (Scheme 1-8).[187,188] The labeling reaction occurs by reaction of cysteine moieties inside the active site with the substrate under release of benzylguanine. It usually proceeds quickly and irreversibly. AGT reacts with a wide variety of substrates (that have an alkylguanine moiety) rendering it a highly attractive labeling method. A possible obstacle is the observed slow but steady unfolding of the protein induced by labeling with certain substrates.

1.2.8.5 Other Labeling Methods for In Vivo *Cell Imaging Applications*

The nonsense suppression technique allows the site-specific introduction of unnatural amino acids into the protein of interest by modified tRNAs.[189] Other labeling methods include peptide labeling based on biotin ligase[190] and Halo-tag (mutant dehalogenase that reacts with certain chloroalkane derivatives).[191] A comprehensive review has been recently published by O'Hare et al.[191]

1.3 Thesis Objective

The objective of this thesis was to develop semi-synthetic proteins for catalytic and analytical applications.

The replacement of organic solvents by environmentally benign solvents such as water is an imperative step towards achieving “green chemistry”. Platinum metals offer powerful catalysts with broad substrate tolerance but are not biologically available. Proteins catalyze a plethora of reactions with high turnover number and high stereoselectivity in their native solvent, which is water. One of the primary objectives of this thesis was to combine the advantages of “both worlds” by generating a semi-synthetic protein. The semi-synthetic protein consists of a catalytically active achiral platinum metal cofactor embedded into a protein framework that induces chirality of the products. The first part of the semi-synthetic proteins for analytical applications section focused on the development and characterization of a synthetic platinum series metal cofactor that is able to catalyze a suitable model reaction in aqueous medium in the absence of the protein. The second part of this topic focused on the catalysis of the model reaction by the synthetic cofactor embedded in a protein.

The exploration of labile metal pools of the biologically relevant transition metals copper, iron and zinc in cells was the purpose of developing semi-synthetic proteins for analytical applications. Fusion proteins consisting of fluorescent proteins and a “reactive” domain that allows attachment of fluorescent or colored chelators (upon metal binding) offer an attractive platform for *in vivo* cell imaging. The genetically encoded fusion proteins are directed to different compartments of interest in the cell. Formation of FRET-pairs upon metal binding allows the detection of the desired transition metals by

changes in the fluorescence emission or lifetime. The exploration of labile metal pools by fluorescence lifetime measurements is an especially promising approach because the fluorescence lifetime is unaffected by the probe concentration or excitation intensity. The focus was on the development and *in vitro* characterization of such semi-synthetic proteins with the goal of reliably measuring the desired labile metal pools *in vivo*.

1.4 References

(1) Collin, R. W. *The Environmental Protection Agency : cleaning up America's act*; Greenwood Press: Westport, Conn., 2006.

(2) Lewis, C. A.; Thunder, J. M. *Federal chemical regulation : TSCA, EPCRA and the Pollution Prevention Act*; Bureau of National Affairs: Washington, D.C., 1997.

(3) Clark, J. H.; Macquarrie, D. J. *Handbook of green chemistry and technology*; Blackwell Science: Oxford [England] ; Malden, MA, 2002.

(4) Anastas, P. T.; Warner, J. C. *Green chemistry : theory and practice*; Oxford University Press: Oxford [England] ; New York, 1998.

(5) Australian Bureau of Statistics.; Australian Bureau of Statistics: Canberra, 1994.

(6) Lancaster, M. *Environment Business Magazine* **2000**, *65*.

(7) *The Nobel Foundation*.

(8) Lindström, U. M. *Organic reactions in water : principles, strategies and applications*; Blackwell Pub.: Oxford ; Ames, Iowa, 2007.

(9) Bruice, P. Y. *Organic chemistry*; 5th ed.; Pearson Education: Upper Saddle River, N.J., 2006.

(10) Furstner, A.; Bogdanovic, B. *Angewandte Chemie-International Edition in English* **1996**, *35*, 2442-2469.

(11) Li, J. J. *Name reactions : a collection of detailed reaction mechanisms*; 3rd expanded ed.; Springer: Berlin ; New York, 2006.

(12) Lindstrom, U. M. *Chemical Reviews* **2002**, *102*, 2751-2771.

(13) Grieco, P. A. *Organic synthesis in water*; 1st ed.; Blackie Academic & Professional: London ; New York, 1998.

(14) Knochel, P.; Augé, J. *Modern solvents in organic synthesis*; Springer: Berlin ; New York, 1999.

(15) Lide, D. R. *Handbook of organic solvents*; CRC Press: Boca Raton, 1995.

(16) Anastas, P. T.; Williamson, T. C. *Green chemistry : frontiers in benign chemical syntheses and processes*; Oxford University Press: Oxford [England] ; New York, 1998.

(17) Dambacher, J.; Zhao, W.; El-Batta, A.; Anness, R.; Jiang, C. C.; Bergdahl, M. *Tetrahedron Letters* **2005**, *46*, 4473-4477.

(18) Itami, K.; Nokami, T.; Yoshida, J. *Advanced Synthesis & Catalysis* **2002**, *344*, 441-451.

(19) Otto, S.; Bertoncin, F.; Engberts, J. *Journal of the American Chemical Society* **1996**, *118*, 7702-7707.

(20) Iwasa, S.; Takezawa, F.; Tuchiya, Y.; Nishiyama, H. *Chemical Communications* **2001**, 59-60.

(21) Breslow, R.; Maitra, U. *Tetrahedron Letters* **1984**, *25*, 1239-1240.

(22) Breslow, R.; Maitra, U.; Rideout, D. *Tetrahedron Letters* **1983**, *24*, 1901-1904.

(23) Rideout, D. C.; Breslow, R. *Journal of the American Chemical Society* **1980**, *102*, 7816-7817.

(24) Grieco, P. A.; Garner, P.; He, Z. *Tetrahedron Letters* **1983**, *24*, 1897-1900.

(25) Grieco, P. A.; Garner, P.; Yoshida, K.; Huffman, J. C. *Tetrahedron Letters* **1983**, *24*, 3807-3810.

(26) Grieco, P. A.; Yoshida, K.; Garner, P. *Journal of Organic Chemistry* **1983**, *48*, 3137-3139.

(27) Blokzijl, W.; Engberts, J. *Angewandte Chemie-International Edition* **1993**, *32*, 1545-1579.

(28) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *Journal of Physical Chemistry B* **2002**, *106*, 521-533.

(29) Chandler, D. *Nature* **2005**, *437*, 640-647.

(30) Pace, C. N.; Shirley, B. A.; McNutt, M.; Gajiwala, K. *Faseb Journal* **1996**, *10*, 75-83.

(31) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*; 5th ed.; W.H. Freeman: New York, 2002.

(32) Breslow, R. *Accounts of Chemical Research* **1991**, *24*, 159-164.

(33) Breslow, R. *Accounts of Chemical Research* **2004**, *37*, 471-478.

(34) Demko, Z. P.; Sharpless, K. B. *Journal of Organic Chemistry* **2001**, *66*, 7945-7950.

(35) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angewandte Chemie-International Edition* **2001**, *40*, 2004-+.

(36) Yamamoto, K. *Advances in organometallic chemistry research*; Nova Science Publishers: New York, 2007.

(37) Barton, D.; Ollis, W. D. *Comprehensive organic chemistry : the synthesis and reactions of organic compounds*; 1st ed.; Pergamon Press: Oxford ; New York, 1979.

(38) Wilkinson, G.; Stone, F. G. A.; Abel, E. W. *Comprehensive organometallic chemistry : the synthesis, reactions, and structures of organometallic compounds*; 1st ed.; Pergamon Press: Oxford [Oxfordshire] ; New York, 1982.

(39) Abel, E. W.; Stone, F. G. A.; Wilkinson, G. *Comprehensive organometallic chemistry II : a review of the literature 1982-1994*; 1st ed.; Pergamon: Oxford ; New York, 1995.

(40) Mingos, D. M. P.; Crabtree, R. H. *Comprehensive organometallic chemistry III*; 1st ed.; Elsevier: Amsterdam ; Boston, 2007.

(41) Cornils, B.; Herrmann, W. A. *Aqueous-phase organometallic catalysis : concepts and applications*; 2nd, completely rev. and enl. ed.; Wiley-VCH: Weinheim, 2004.

(42) Weissermel, K.; Arpe, H.-J. *Industrial organic chemistry*; 4th completely rev. ed.; Wiley-VCH: Weinheim, 2003.

(43) Campos-Malpartida, T.; Fekete, M.; Joo, F.; Katho, A.; Romerosa, A.; Saoud, M.; Wojtkow, W. *Journal of Organometallic Chemistry* **2008**, *693*, 468-474.

(44) Chen, L.; Li, C. J. *Advanced Synthesis & Catalysis* **2006**, *348*, 1459-1484.

(45) Eckl, R. W.; Priermeier, T.; Herrmann, W. A. *Journal of Organometallic Chemistry* **1997**, *532*, 243-249.

(46) Herrerias, C. I.; Yao, X. Q.; Li, Z. P.; Li, C. J. *Chemical Reviews* **2007**, *107*, 2546-2562.

(47) Joo, F. *Accounts of Chemical Research* **2002**, *35*, 738-745.

(48) Joo, F.; Katho, A. *Journal of Molecular Catalysis a-Chemical* **1997**, *116*, 3-26.

(49) Kovacs, G.; Ujaque, G.; Lledos, A.; Joo, F. *European Journal of Inorganic Chemistry* **2007**, 2879-2889.

(50) Li, C. J. *Accounts of Chemical Research* **2002**, *35*, 533-538.

(51) Li, C. J.; Chen, L. *Chemical Society Reviews* **2006**, *35*, 68-82.

(52) Rossin, A.; Kovacs, G.; Ujaque, G.; Lledos, A.; Joo, F. *Organometallics* **2006**, *25*, 5010-5023.

(53) Li, C. J. *Chemical Reviews* **2005**, *105*, 3095-3165.

(54) Joó, F. *Aqueous organometallic catalysis*; Kluwer Academic Publishers: Dordrecht ; Boston, 2001.

(55) Cornils, B.; Kuntz, E. G. *Journal of Organometallic Chemistry* **1995**, *502*, 177-186.

(56) Brynner, R.; Stephens, T. D. *Dark remedy : the impact of Thalidomide and its revival as a vital medicine*; Basic Books: New York, 2001.

(57) Pan, C. F.; Wang, Z. Y. *Coordination Chemistry Reviews* **2008**, *252*, 736-750.

(58) Ahlford, K.; Lind, J.; Maler, L.; Adolfsson, H. *Green Chemistry* **2008**, *10*, 832-835.

(59) Noyori, R.; Hashiguchi, S. *Accounts of Chemical Research* **1997**, *30*, 97-102.

(60) Jacobsen, E. N.; Pfaltz, A.; Yamamoto, H. *Comprehensive asymmetric catalysis*; Springer: Berlin ; New York, 1999.

(61) Jacobsen, E. N.; Pfaltz, A.; Yamamoto, H. *Comprehensive asymmetric catalysis. Supplement*; Springer: Berlin ; New York, 2004.

(62) Bairoch, A. *Nucleic Acids Research* **2000**, *28*, 304-305.

(63) Davis, B. G.; Borer, V. *Natural Product Reports* **2001**, *18*, 618-640.

(64) Garcia-Junceda, E.; Garcia-Garcia, J. F.; Bastida, A.; Fernandez-Mayoralas, A. *Bioorganic & Medicinal Chemistry* **2004**, *12*, 1817-1834.

(65) Johnson, C. R. *Accounts of Chemical Research* **1998**, *31*, 333-341.

(66) Margolin, A. L. *Enzyme and Microbial Technology* **1993**, *15*, 266-280.

(67) Panke, S.; Wubbolts, M. *Current Opinion in Chemical Biology* **2005**, *9*, 188-194.

(68) Patel, R. N. In *Advances in Applied Microbiology, Vol 47* 2000; Vol. 47, p 33-78.

(69) Zaks, A.; Dodds, D. R. *Drug Discovery Today* **1997**, *2*, 513-531.

(70) Olpe, H. R.; Demieville, H.; Baltzer, V.; Bencze, W. L.; Koella, W. P.; Wolf, P.; Haas, H. L. *European Journal of Pharmacology* **1978**, *52*, 133-136.

(71) Chenevert, R.; Desjardins, M. *Tetrahedron Letters* **1991**, *32*, 4249-4250.

(72) Drauz, K.; Waldmann, H. *Enzyme catalysis in organic synthesis : a comprehensive handbook*; 2nd, completely rev. and enl. ed.; Wiley-VCH: Weinheim ; New York, 2002.

(73) Dolphin, D. *The Porphyrins*; Academic Press: New York, 1978.

(74) Harada, K.; Makino, M.; Sugimoto, H.; Hirota, S.; Matsuo, T.; Shiro, Y.; Hisaeda, Y.; Hayashi, T. *Biochemistry* **2007**, *46*, 9406-9416.

(75) Hayashi, T.; Dejima, H.; Matsuo, T.; Sato, H.; Murata, D.; Hisaeda, Y. *Journal of the American Chemical Society* **2002**, *124*, 11226-11227.

(76) Hayashi, T.; Hisaeda, Y. *Accounts of Chemical Research* **2002**, *35*, 35-43.

(77) Hayashi, T.; Murata, D.; Makino, M.; Sugimoto, H.; Matsuo, T.; Sato, H.; Shiro, Y.; Hisaeda, Y. *Inorganic Chemistry* **2006**, *45*, 10530-10536.

(78) Matsuo, T.; Dejima, H.; Hirota, S.; Murata, D.; Sato, H.; Ikegami, T.; Hori, H.; Hisaeda, Y.; Hayashi, T. *Journal of the American Chemical Society* **2004**, *126*, 16007-16017.

(79) Matsuo, T.; Ikegami, T.; Sato, H.; Hisaeda, Y.; Hayashi, T. *Journal of Inorganic Biochemistry* **2006**, *100*, 1265-1271.

(80) Matsuo, T.; Ito, K.; Nakashima, Y.; Hisaeda, Y.; Hayashi, T. *Journal of Inorganic Biochemistry* **2008**, *102*, 166-173.

(81) Matsuo, T.; Murata, D.; Hisaeda, Y.; Hori, H.; Hayashi, T. *Journal of the American Chemical Society* **2007**, *129*, 12906-+.

(82) Matsuo, T.; Nagai, H.; Hisaeda, Y.; Hayashi, T. *Chemical Communications* **2006**, 3131-3133.

(83) Matsuo, T.; Tsuruta, T.; Maehara, K.; Sato, H.; Hisaeda, Y.; Hayashi, T. *Inorganic Chemistry* **2005**, *44*, 9391-9396.

(84) Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z. L.; Guo, Z. J.; Lu, Y. *Journal of the American Chemical Society* **2004**, *126*, 10812-10813.

(85) Lu, Y. *Current Opinion in Chemical Biology* **2005**, *9*, 118-126.

(86) Lu, Y. *Inorganic Chemistry* **2006**, *45*, 9930-9940.

(87) Lu, Y. *Angewandte Chemie-International Edition* **2006**, *45*, 5588-5601.

(88) Lu, Y.; Berry, S. M.; Pfister, T. D. *Chemical Reviews* **2001**, *101*, 3047-3080.

(89) Lu, Y.; Valentine, J. S. *Current Opinion in Structural Biology* **1997**, *7*, 495-500.

(90) Sigman, J. A.; Kim, H. K.; Zhao, X. A.; Carey, J. R.; Lu, Y. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 3629-3634.

(91) Sigman, J. A.; Kwok, B. C.; Lu, Y. *Journal of the American Chemical Society* **2000**, *122*, 8192-8196.

(92) Zhang, J. L.; Garner, D. K.; Liang, L.; Chen, Q.; Lu, Y. *Chemical Communications* **2008**, 1665-1667.

(93) Zhao, X.; Nilges, M. J.; Lu, Y. *Biochemistry* **2005**, *44*, 6559-6564.

(94) Zhao, X.; Yeung, N.; Russell, B. S.; Garner, D. K.; Lu, Y. *Journal of the American Chemical Society* **2006**, *128*, 6766-6767.

(95) Zhao, X.; Yeung, N.; Wang, Z. L.; Guo, Z. J.; Lu, Y. *Biochemistry* **2005**, *44*, 1210-1214.

(96) Adachi, S.; Nagano, S.; Ishimori, K.; Watanabe, Y.; Morishima, I.; Egawa, T.; Kitagawa, T.; Makino, R. *Biochemistry* **1993**, *32*, 241-252.

(97) Hara, I.; Ueno, T.; Ozaki, S.; Itoh, S.; Lee, K.; Ueyama, N.; Watanabe, Y. *Journal of Biological Chemistry* **2001**, *276*, 36067-36070.

(98) Kato, S.; Yang, H. J.; Ueno, T.; Ozaki, S.; Phillips, G. N.; Fukuzumi, S.; Watanabe, Y. *Journal of the American Chemical Society* **2002**, *124*, 8506-8507.

(99) Matsui, T.; Ozaki, S.; Liong, E.; Phillips, G. N.; Watanabe, Y. *Journal of Biological Chemistry* **1999**, *274*, 2838-2844.

(100) Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angewandte Chemie-International Edition* **2003**, *42*, 1005-+.

(101) Ozaki, S.; Hara, I.; Matsui, T.; Watanabe, Y. *Biochemistry* **2001**, *40*, 1044-1052.

(102) Ozaki, S.; Matsui, T.; Roach, M. P.; Watanabe, Y. *Coordination Chemistry Reviews* **2000**, *198*, 39-59.

(103) Ozaki, S.; Yang, H. J.; Matsui, T.; Goto, Y.; Watanabe, Y. *Tetrahedron-Asymmetry* **1999**, *10*, 183-192.

(104) Ozaki, S. I.; Roach, M. P.; Matsui, T.; Watanabe, Y. *Accounts of Chemical Research* **2001**, *34*, 818-825.

(105) Roach, M. P.; Ozaki, S.; Watanabe, Y. *Biochemistry* **2000**, *39*, 1446-1454.

(106) Roach, M. P.; Puspita, W. J.; Watanabe, Y. *Journal of Inorganic Biochemistry* **2000**, *81*, 173-182.

(107) Sato, H.; Hayashi, T.; Ando, T.; Hisaeda, Y.; Ueno, T.; Watanabe, Y. *Journal of the American Chemical Society* **2004**, *126*, 436-437.

(108) Ueno, T.; Koshiyama, T.; Abe, S.; Yokoi, N.; Ohashi, M.; Nakajima, H.; Watanabe, Y. *Journal of Organometallic Chemistry* **2007**, *692*, 142-147.

(109) Ueno, T.; Koshiyama, T.; Ohashi, M.; Kondo, K.; Kono, M.; Suzuki, A.; Yamane, T.; Watanabe, Y. *Journal of the American Chemical Society* **2005**, *127*, 6556-6562.

(110) Ueno, T.; Ohashi, M.; Kono, M.; Kondo, K.; Suzuki, A.; Yamane, T.; Watanabe, Y. *Inorganic Chemistry* **2004**, *43*, 2852-2858.

(111) Watanabe, Y. *Reviews on Heteroatom Chemistry* **2000**, *22*, 135-150.

(112) Watanabe, Y. *Current Opinion in Chemical Biology* **2002**, *6*, 208-216.

(113) Yang, H. J.; Matsui, T.; Ozaki, S.; Kato, S.; Ueno, T.; Phillips, G. N.; Fukuzumi, S.; Watanabe, Y. *Biochemistry* **2003**, *42*, 10174-10181.

(114) Tominaga, M.; Kumagai, T.; Takita, S.; Taniguchi, I. *Chemistry Letters* **1993**, 1771-1774.

(115) Phillips, S. E. V. *Journal of Molecular Biology* **1980**, *142*, 531-554.

(116) Hamachi, I.; Shinkai, S. *European Journal of Organic Chemistry* **1999**, 539-549.

(117) Wilson, M. E.; Whitesides, G. M. *Journal of the American Chemical Society* **1978**, *100*, 306-307.

(118) Steinreiber, J.; Ward, T. R. *Coordination Chemistry Reviews* **2008**, *252*, 751-766.

(119) Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *Journal of the American Chemical Society* **2004**, *126*, 14411-14418.

(120) Klein, G.; Humbert, N.; Gradinaru, J.; Ivanova, A.; Gilardoni, F.; Rusbandi, U. E.; Ward, T. R. *Angewandte Chemie-International Edition* **2005**, *44*, 7764-7767.

(121) Pordea, A.; Creus, M.; Panek, J.; Duboc, C.; Mathis, D.; Novic, M.; Ward, T. R. *Journal of the American Chemical Society* **2008**, *130*, 8085-8088.

(122) Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. *Angewandte Chemie-International Edition* **2008**, *47*, 701-705.

(123) Colonna, S.; Manfredi, A.; Annunziata, R. *Tetrahedron Letters* **1988**, *29*, 3347-3350.

(124) Colonna, S.; Gaggero, N.; Manfredi, A.; Spadoni, M.; Casella, L.; Carrea, G.; Pasta, P. *Tetrahedron* **1988**, *44*, 5169-5178.

(125) Colonna, S.; Manfredi, A.; Annunziata, R.; Spadoni, M. *Tetrahedron* **1987**, *43*, 2157-2164.

(126) Colonna, S.; Manfredi, A. *Tetrahedron Letters* **1986**, *27*, 387-390.

(127) Sugimoto, T.; Kokubo, T.; Miyazaki, J.; Tanimoto, S.; Okano, M. *Journal of the Chemical Society-Chemical Communications* **1979**, 402-404.

(128) Baba, N.; Matsumura, Y.; Sugimoto, T. *Tetrahedron Letters* **1978**, 4281-4284.

(129) Mahammed, A.; Gross, Z. *Journal of the American Chemical Society* **2005**, *127*, 2883-2887.

(130) Mahammed, A.; Gray, H. B.; Weaver, J. J.; Sorasaenee, K.; Gross, Z. *Bioconjugate Chemistry* **2004**, *15*, 738-746.

(131) Reetz, M. T.; Jiao, N. *Angewandte Chemie-International Edition* **2006**, *45*, 2416-2419.

(132) Davies, R. R.; Distefano, M. D. *Journal of the American Chemical Society* **1997**, *119*, 11643-11652.

(133) Reetz, M. T. *Tetrahedron* **2002**, *58*, 6595-6602.

(134) Roy, R. S.; Imperiali, B. *Protein Engineering* **1997**, *10*, 691-698.

(135) Silva, J. J. R. F. d.; Williams, R. J. P. *The biological chemistry of the elements : the inorganic chemistry of life*; 2nd ed.; Oxford University Press: Oxford ; New York, 2001.

(136) Barceloux, D. G. *Journal of Toxicology-Clinical Toxicology* **1999**, *37*, 217-230.

(137) Tapiero, H.; Townsend, D. M.; Tew, K. D. *Biomedicine & Pharmacotherapy* **2003**, *57*, 386-398.

(138) Crichton, R. R.; Boelaert, J. R. *Inorganic biochemistry of iron metabolism : from molecular mechanisms to clinical consequences*; 2nd ed.; Wiley: Chichester ; New York, 2001.

(139) Vallee, B. L.; Falchuk, K. H. *Physiological Reviews* **1993**, *73*, 79-118.

(140) Fabris, N.; Mocchegiani, E. *Aging-Clinical and Experimental Research* **1995**, *7*, 77-93.

(141) Hershfinkel, M.; Silverman, W. F.; Sekler, I. 2007, p 331-336.

(142) Apgar, J. *Annual Review of Nutrition* **1985**, *5*, 43-68.

(143) Bedwal, R. S.; Bahuguna, A. *Experientia* **1994**, *50*, 626-640.

(144) Takeda, A. *Biometals* **2001**, *14*, 343-351.

(145) Ibs, K. H.; Rink, L. *Journal of Nutrition* **2003**, *133*, 1452S-1456S.

(146) Salgueiro, M. J.; Zubillaga, M.; Lysionek, A.; Cremaschi, G.; Goldman, C. G.; Caro, R.; De Paoli, T.; Hager, A.; Weill, R.; Boccio, J. *Biological Trace Element Research* **2000**, *76*, 193-205.

(147) Solomons, N. W. *Nutrition Reviews* **1998**, *56*, 27-28.

(148) Aisen, P.; Wessling-Resnick, M.; Leibold, E. A. *Current Opinion in Chemical Biology* **1999**, *3*, 200-206.

(149) Finney, L. A.; O'Halloran, T. V. *Science* **2003**, *300*, 931-936.

(150) Harrison, M. D.; Dameron, C. T. *Journal of Biochemical and Molecular Toxicology* **1999**, *13*, 93-106.

(151) Brissot, P.; Troadec, M. B.; Bardou-Jacquet, E.; Le Lan, C.; Jouanolle, A. M.; Deugnier, Y.; Loreal, O. *Blood Reviews* **2008**, *22*, 195-210.

(152) Herd, S. M.; Camakaris, J.; Christofferson, R.; Wookey, P.; Danks, D. M. *Biochemical Journal* **1987**, *247*, 341-347.

(153) Williams, R. J. P.; da Silva, J. *Coordination Chemistry Reviews* **2000**, *200*, 247-348.

(154) Breuer, W.; Shvartsman, M.; Cabantchik, Z. I. *International Journal of Biochemistry & Cell Biology* **2008**, *40*, 350-354.

(155) Maret, W.; Larsen, K. S.; Vallee, B. L. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 2233-2237.

(156) Esposito, B. P.; Epsztejn, S.; Breuer, W.; Cabantchik, Z. I. *Analytical Biochemistry* **2002**, *304*, 1-18.

(157) Haase, H.; Beyersmann, D. *Biometals* **1999**, *12*, 247-254.

(158) Suhy, D. A.; Ohalloran, T. V. In *Metal Ions in Biological Systems, Vol 32* 1996; Vol. 32, p 557-578.

(159) Grynkiewicz, G.; Poenie, M.; Tsien, R. Y. *Journal of Biological Chemistry* **1985**, *260*, 3440-3450.

(160) Lakowicz, J. R. *Principles of fluorescence spectroscopy*; 3rd ed.; Springer: New York, 2006.

(161) Lytton, S. D.; Mester, B.; Libman, J.; Shanzer, A.; Cabantchik, Z. I. *Analytical Biochemistry* **1992**, *205*, 326-333.

(162) Werts, M. H. V.; Hofstraat, J. W.; Geurts, F. A. J.; Verhoeven, J. W. *Chemical Physics Letters* **1997**, *276*, 196-201.

(163) Mahadevan, I. B.; Kimber, M. C.; Lincoln, S. F.; Tiekink, E. R. T.; Ward, A. D.; Betts, W. H.; Forbes, I. J.; Zalewski, P. D. *Australian Journal of Chemistry* **1996**, *49*, 561-568.

(164) Hirano, T.; Kikuchi, K.; Urano, Y.; Nagano, T. *Journal of the American Chemical Society* **2002**, *124*, 6555-6562.

(165) Zeng, L.; Miller, E. W.; Pralle, A.; Isacoff, E. Y.; Chang, C. J. *Journal of the American Chemical Society* **2006**, *128*, 10-11.

(166) Yang, L. C.; McRae, R.; Henary, M. M.; Patel, R.; Lai, B.; Vogt, S.; Fahrni, C. J. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 11179-11184.

(167) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006**, *312*, 217-224.

(168) Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. *Nature Methods* **2005**, *2*, 905-909.

(169) http://nobelprize.org/nobel_prizes/chemistry/laureates/2008/ Jan-2009

(170) Dansen, T. B.; Wirtz, K. W. A.; Wanders, R. J. A.; Pap, E. H. W. *Nature Cell Biology* **2000**, *2*, 51-53.

(171) Hicks, G. R.; Raikhel, N. V. *Annual Review of Cell and Developmental Biology* **1995**, *11*, 155-188.

(172) Nilsson, T.; Warren, G. *Current Opinion in Cell Biology* **1994**, *6*, 517-521.

(173) Rizzuto, R.; Brini, M.; Pizzo, P.; Murgia, M.; Pozzan, T. *Current Biology* **1995**, *5*, 635-642.

(174) Piston, D. W.; Kremers, G. J. *Trends in Biochemical Sciences* **2007**, *32*, 407-414.

(175) Wallrabe, H.; Periasamy, A. *Current Opinion in Biotechnology* **2005**, *16*, 19-27.

(176) Hahn, M. E.; Muir, T. W. *Trends in Biochemical Sciences* **2005**, *30*, 26-34.

(177) Yeo, D. S. Y.; Srinivasan, R.; Chen, G. Y. J.; Yao, S. Q. *Chemistry-a European Journal* **2004**, *10*, 4664-4672.

(178) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. *Journal of the American Chemical Society* **2002**, *124*, 6063-6076.

(179) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269-272.

(180) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nature Reviews Molecular Cell Biology* **2002**, *3*, 906-918.

(181) Paulus, H. *Annual Review of Biochemistry* **2000**, *69*, 447-496.

(182) Noren, C. J.; Wang, J. M.; Perler, F. B. *Angewandte Chemie-International Edition* **2000**, *39*, 450-466.

(183) Ludwig, C.; Pfeiff, M.; Linne, U.; Mootz, H. D. *Angewandte Chemie-International Edition* **2006**, *45*, 5218-5221.

(184) Sun, W. C.; Yang, J.; Liu, X. Q. *Journal of Biological Chemistry* **2004**, *279*, 35281-35286.

(185) Muir, T. W. *Annual Review of Biochemistry* **2003**, *72*, 249-289.

(186) Lue, R. Y. P.; Chen, G. Y. J.; Hu, Y.; Zhu, Q.; Yao, S. Q. *Journal of the American Chemical Society* **2004**, *126*, 1055-1062.

(187) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nature Biotechnology* **2003**, *21*, 86-89.

(188) Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 9955-9959.

(189) Wang, L.; Schultz, P. G. *Chemical Communications* **2002**, 1-11.

(190) Chen, I.; Ting, A. Y. *Current Opinion in Biotechnology* **2005**, *16*, 35-40.

(191) O'Hare, H. M.; Johnsson, K.; Gautier, A. *Current Opinion in Structural Biology* **2007**, *17*, 488-494.

Chapter II

CHARACTERIZATION OF RUTHENIUM PORPHYRINS AS WATER-SOLUBLE ORGANOMETALLIC CATALYSTS

2.1 Introduction and Concept

The replacement of organic solvents by environmentally benign solvents such as water is an imperative step towards achieving "green chemistry" (see chapter 1.1.1).[1] Utilizing semi-synthetic proteins for catalysis with a cofactor containing a platinum metal requires either thermodynamic or kinetic stability of the platinum metal catalyst and all reactants in the native solvent of the protein, which is water. Many organometallic catalysts are known to undergo reactions in water, which lead to their deactivation (see chapter 1.1.2).[2] Therefore, the focus of this study was to explore, whether catalysis with a platinum metal porphyrin (without being attached to a protein) can be carried out successfully in aqueous medium. Properties such as turnover numbers, kinetics and the influence of axial ligands on the reactivity of the catalyst were studied. These characteristics are important for the utility of the catalyst, embedded in a protein.

2.1.1 Semi-Synthetic Protein Designs Based on Incorporation of Cofactors

In order to choose a platinum metal cofactor that can be incorporated into a protein and exhibit catalytic activity, the methods of embedding a cofactor into a protein have to be considered. The incorporation of an organometallic cofactor into a protein can in principle be accomplished by three basic approaches.[3,4]

First, an appropriate metal binding site can be directly generated in the active site of the protein by site-directed mutagenesis utilizing naturally occurring amino acids as coordinating ligands. This type of design requires a very detailed understanding of the coordination chemistry of amino acid ligands bound to the non-biological metal center. Moreover, any influences of the metal on the conformation of the active site or the surrounding protein backbone upon incorporation are difficult to predict. Thus, simple reconstitution of a metalloprotein with a platinum series metal would probably yield a catalytically inactive enzyme.

The second method to create a semi-synthetic protein containing a platinum metal series organometallic cofactor is by supramolecular attachment. This highly promising approach was greatly optimized for hydrogenation reactions with rhodium based catalysts, based on the streptavidin/avidin biotin technology,[5] by Ward and coworkers;[6-14] however, this viable design of artificial metalloenzymes utilizes poorly defined structures. Therefore supramolecular anchoring attempts by the streptavidin/avidin biotin non-covalent interactions were not pursued.

Finally, an organometallic cofactor can be attached to a protein framework by the cofactor reconstitution method.[15] In this method the metal cation is embedded as a prebuilt synthetic prosthetic group. The incorporation of either natural or modified

cofactors by this method was carried out successfully for several hemoproteins and flavoenzymes.[16-18] An important benefit of this approach is that the organometallic cofactor is already catalytically active, even if it is not incorporated in a protein. Thus, the synthetic cofactor can already be tested and optimized for reactions in aqueous media. Moreover, the cofactor reconstitution method allows the incorporation of the cofactor into the structurally clearly defined (if structural information is available) active site. The active site pocket provides a chiral framework that induces stereoselectivity. Reaction properties of the resulting semi-synthetic enzyme, including turnover number, substrate selectivity and enantioselectivity, can be enhanced by site-directed mutagenesis of crucial active-site residues.

Heme proteins and enzymes contain heme groups, which are metalloporphyrins (usually Fe-porphyrins). Metalloporphyrins are ubiquitous organometallic cofactors that are involved in the binding of oxygen (e.g. hemoglobin and myoglobin) and catalyze a plethora of reactions, such as oxidations (e.g. peroxidases and cytochrome c oxidase) and electron transfer reactions (e.g. cytochrome c).[16,19,20] The porphyrin prosthetic groups are mostly bound in the active site by hydrophobic interactions and hydrogen bonding. Due to the lack of covalent interactions they often can be readily extracted at acidic pH.[15] Despite some losses by denaturation of the protein due to the low pH, the reconstitution of the protein with the cofactor usually leads to the recovery of the original enzyme activity. Consequently, the reconstitution of the apoprotein with a synthetically modified metalloporphyrin cofactor, containing a biologically not available platinum series metal, represented a promising strategy that was pursued (see chapter 3).

2.1.2 Selection of Platinum Metal in Cofactor

Ruthenium organometallic complexes facilitate the catalysis of myriad reactions.[21] Ruthenium porphyrins catalyze the cyclopropanation of a variety of olefins, [22-32] thus rendering this type of reaction an attractive starting point for the development of a catalytically active semi-synthetic protein. While iron porphyrins are efficient catalysts for the cyclopropanation of olefins this prosthetic group was originally evolved to bind dioxygen and would primarily function as oxidation catalyst or oxygen carrier.[33,34] Many peroxidases, containing iron porphyrins as cofactor, can facilitate the enantioselective epoxidation of a broad range of olefins utilizing hydrogen peroxide as oxidizing agent.[35-37]

2.1.3 Selection of Model Reaction

Epoxidation and cyclopropanation are closely related organic transformations of olefins (Figure 2-1). In case of the epoxidation reaction, an electrophilic iron-oxo or peroxo porphyrin species is attacked by the electron rich C=C double bond of the olefin to give an epoxide. In the cyclopropanation reaction a Fisher-type ruthenium porphyrin carbene intermediate is formed, which is electrophilic and reacts with the electron rich alkene double bond yielding a cyclopropane species. The carbene complex is only slightly larger than the iron-oxo intermediate rendering the active site of hemoproteins, capable of catalyzing epoxidation reactions, spacious enough to accommodate the carbene intermediates of the cyclopropanation reaction.

Cyclopropanes are attractive synthetic targets because the reactivity of the three-membered ring due to their high ring-strain renders them useful as synthetic building

blocks. Cyclopropane derivatives can be found in numerous natural products and biologically active compounds.

A key requirement for the creation of a “cyclopropanase” is that the cofactor is catalytically active under the conditions required by its protein-host. Since proteins require water as their natural solvent the ruthenium porphyrin catalyst, its carbene intermediate and all involved reagents have to be stable enough in aqueous medium to successfully catalyze the cyclopropanation reaction (Scheme 2-7).

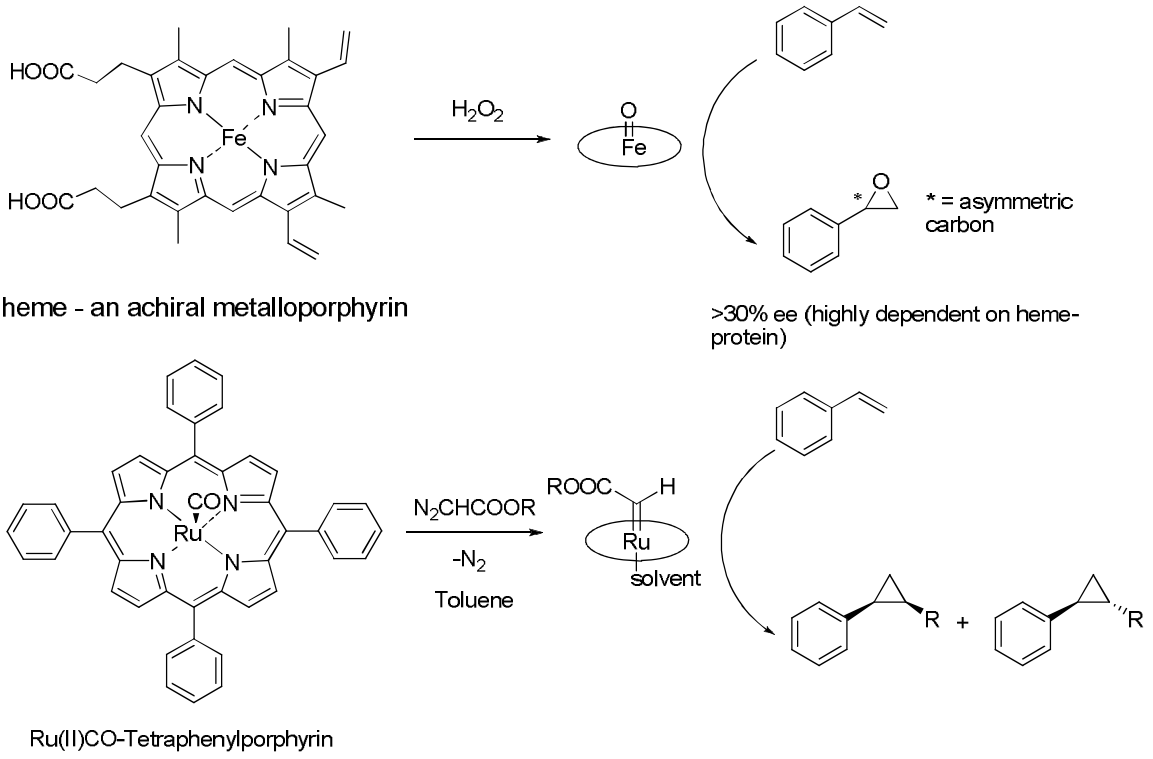
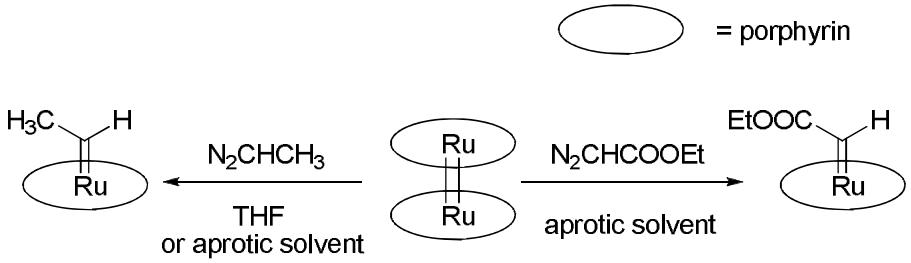


Figure 2-1. Heme catalyzed peroxidation and ruthenium porphyrin catalyzed cyclopropanation reaction.

2.1.4 Selection of a Suitable Ruthenium Porphyrin for Aqueous Cyclopropanation Model Reactions

In order to successfully carry out cyclopropanation reactions in aqueous media the catalyst needs to be water-soluble (in case of homogenous catalysis). The solubilization of chemical compounds for applications in water can be achieved by adding polar functional groups, such as the alcohol, amine, amide, carboxyl, phosphate and sulfonate groups.[38] Moreover, the ruthenium porphyrin should contain bulky moieties to inhibit or reduce dimerization. Dimerization is a known feature of certain ruthenium porphyrins and can still result in catalytically active species (Figure 2-2).[39] Due to hydrophobic effects, dimerization of the ruthenium catalyst might even be more favorable in water than in organic solvents. However, a dimer is not representative as a model for possible reactions with reconstituted proteins because the active site pocket will most likely not be large enough to fit a dimer. Besides, the ruthenium porphyrin should have a relatively simple structure that does not require a laborious multi-step synthesis. A symmetric structure simplifies spectral studies, which is an important feature for mechanistic studies. Moreover, the synthesis should lead to sufficient quantities for detailed analytical studies.

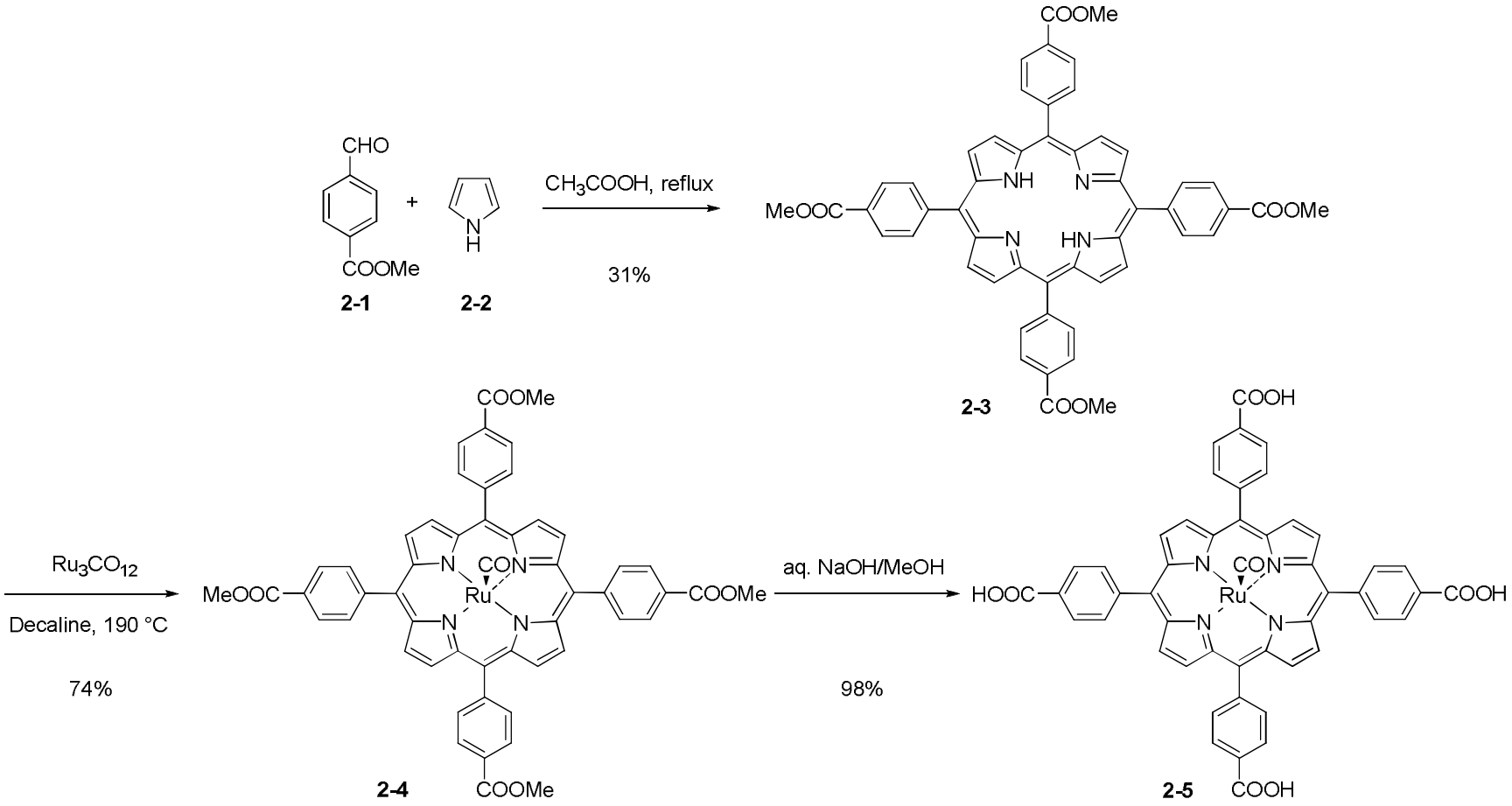
Ruthenium(II)carbonyl *meso*-tetrakis-(4-carboxyphenyl)porphyrin (Ru(CO)TPP-acid) encompasses all these requirements (Scheme 2-2). It has four carboxylic acids to increase the water-solubility, is highly symmetrical and can be synthesized in suitable quantities in three steps.



Scheme 2-1.[39]

2.2 Synthesis

The synthesis of Ru(CO)TPP-acid commenced with the acid catalyzed reaction of the two commercially available starting materials methyl 4-formylbenzoate (**2-1**) and pyrrole (**2-2**), yielding *meso*-tetrakis-(4-carboxyphenyl)porphyrin tetramethyl ester (**2-3**).[40] The ruthenium carbonyl moiety was inserted into **2-3** by reaction with $Ru_3(CO)_{12}$ at high temperatures (190 ℃) leading to ruthenium(II) carbonyl *meso*-tetrakis-(4-carboxyphenyl)porphyrin tetramethyl ester (**2-4**).[41] The water-soluble final ruthenium porphyrin **2-5** was obtained by hydrolysis with NaOH in aqueous methanol.[42]


CHO
COOMe
2-1
2-2
CH3COOH, reflux
31%
COOMe
MeOOC
COOMe
NH
N
N
HN
COOMe
2-3
Ru3CO12
Decaline, 190 °C
74%
COOMe
MeOOC
COOMe
Ru
CO
2-4
COOMe
aq. NaOH/MeOH
98%
COOH
HOOC
COOH
Ru
CO
2-5
COOH

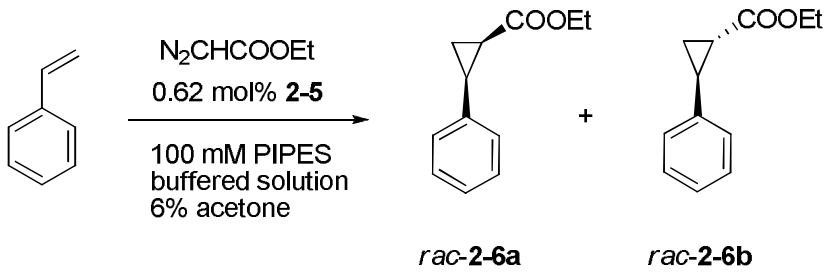
Scheme 2-2.

2.3 Model Cyclopropanation Reactions in Aqueous Medium

2.3.1 Selection of Reaction Conditions

In choosing suitable reaction conditions, the most important parameters are the type of diazo compound acting as a "carbene-donor", the olefin, buffer, pH and a cosolvent. Cyclopropanation reactions with ruthenium organometallic catalysts are most commonly carried out with either methyl diazoacetate or ethyl diazoacetate due to their commercial availability and relative stability (explosive only at elevated temperatures unlike diazomethane).[43] Styrene is the simplest aromatic terminal alkene and most widely used for cyclopropanation reactions despite its tendency to polymerize. Styrene's lipophilicity and subsequent insolubility in water can be overcome by the addition of a water-miscible cosolvent, such as acetone or acetonitrile. Common buffers such as phosphate and acetate buffers often are not suitable for organometallic transformations because the formed ions can bind to metal center and subsequently change the reactivity. An alternative non-coordinating buffer is 1,4-piperazinediethanesulfonic acid (PIPES), suitable for buffering around neutral the pH (pK_a = 7.2).

Hence, the model cyclopropanation reactions were conducted with **2-5**, EDA and styrene in a PIPES buffered solution with acetone as cosolvent (Scheme 2-3).



Scheme 2-3.

2.3.2 Analysis of the Model Cyclopropanation by GC-MS

The cyclopropanation of styrene by slow addition of ethyl diazoacetate was catalyzed by Ru(CO)TPP-acid (**2-5**) in aqueous buffered solution at neutral pH (Scheme 2-3). The reaction mixture was analyzed by coupled gas chromatography – mass spectrometry (GC-MS). Four compounds could be reliable identified (other compounds were either solvent additives or known impurities in the column): styrene, diethyl maleate, *cis*-1-carbethoxy-2-phenylcyclopropane (**2-6a**) and *trans*-1-carbethoxy-2-phenylcyclopropane (**2-6b**) (Figure 2-2). The absence of the starting material ethyl diazoacetate despite unreacted styrene is either caused by decomposition during the reaction or by thermal decomposition in the GC column (due to the gas chromatography process), which is a well established property of diazo-compounds.[44] The formation of diethyl maleate and diethyl fumarate is a common side-reaction for ruthenium-based catalysis with EDA.[43] In fact, ruthenium porphyrins that are similar to **2-5** are efficient catalysts for the formation of these molecules in the absence of styrene (Scheme 2-4).[26]

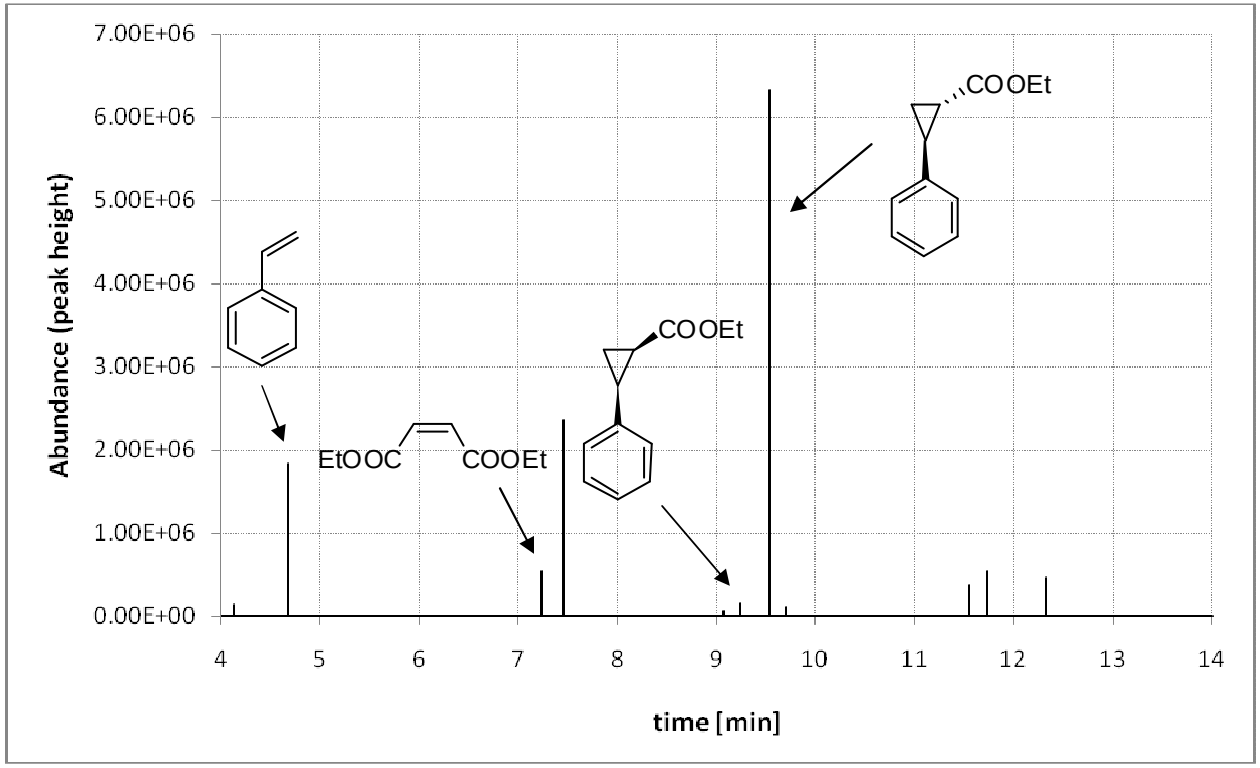
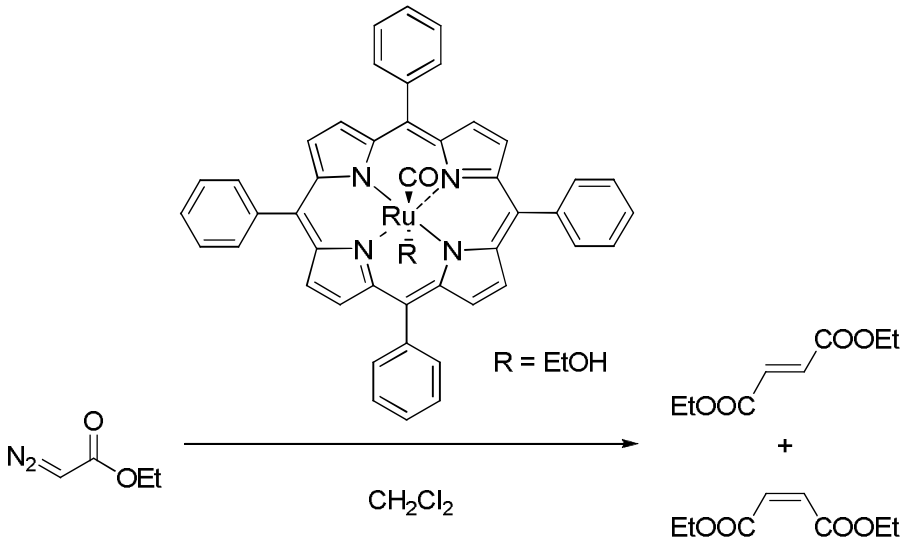


Figure 2-2. GC-MS chromatogram of cyclopropanation reaction. An HP-5 column was used ((5%-Phenyl)-methylpolysiloxane); temperature ramp 50 ℃ to 200 ℃ (4 minutes; 15 minutes total run time).

Interestingly, the formation of the *trans*-sideproduct diethyl fumarate was not observed. The pathway for its formation is either not accessible in water or the small amount of fumarate, formed during the cyclopropanation reaction, is below the detection limit of the instrument.



Scheme 2-4.

2.3.3 Quantitation by GC-FID

Several requirements have to be met in order to quantify the compound compostion accurately by GC-FID (FID = Flame Ionization Detection). The crucial requirement of volatility and mobility in the stationary phase has already been met (Chapter 2.3.1). Since it is desirable to use an internal standard, such as tetradecane or dodecane, to correct for the loss of analyte during sample preparation or sample inlet the measured ratios of analyte to internal standard have to be linear in the desired range. Therefore a series of solutions with varying analyte concentrations of styrene, diethyl

maleate and the trans-cyclopropanation product **2-6b** with constant amounts of dodecane was prepared and analyzed by GC-FID (Figure 2-4).

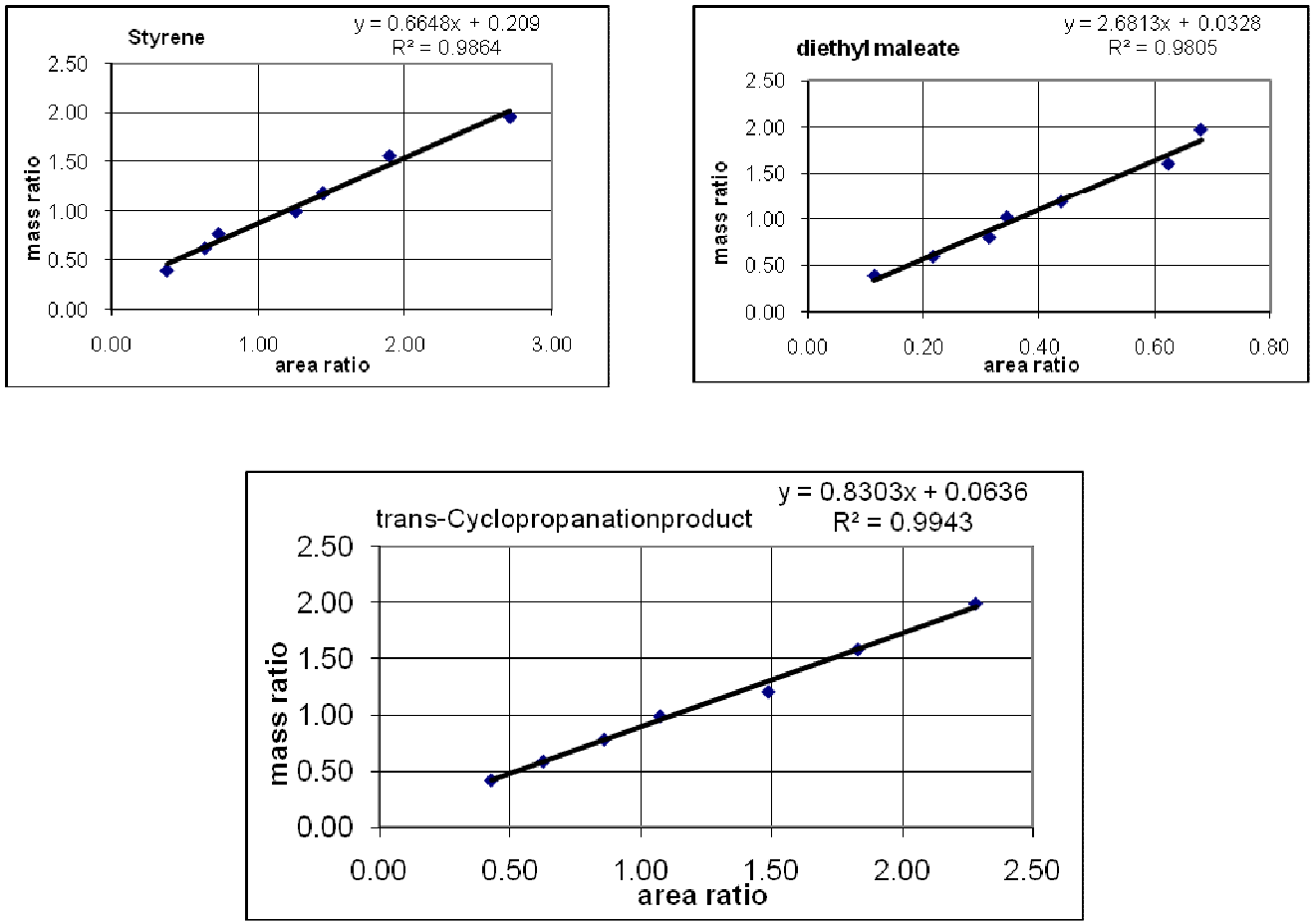


Figure 2-4. Standardization graphs of styrene, diethyl maleate and **2-6b**.

Plotting of mass analyte to mass internal standard (dodecane) ratio against peak area analyte to peak area internal standard ratio provided standardization graphs. These showed exclusively linear behavior in the desired concentration region. Thus, all relevant analytes could be reliably quantified by GC-FID in the further studies.

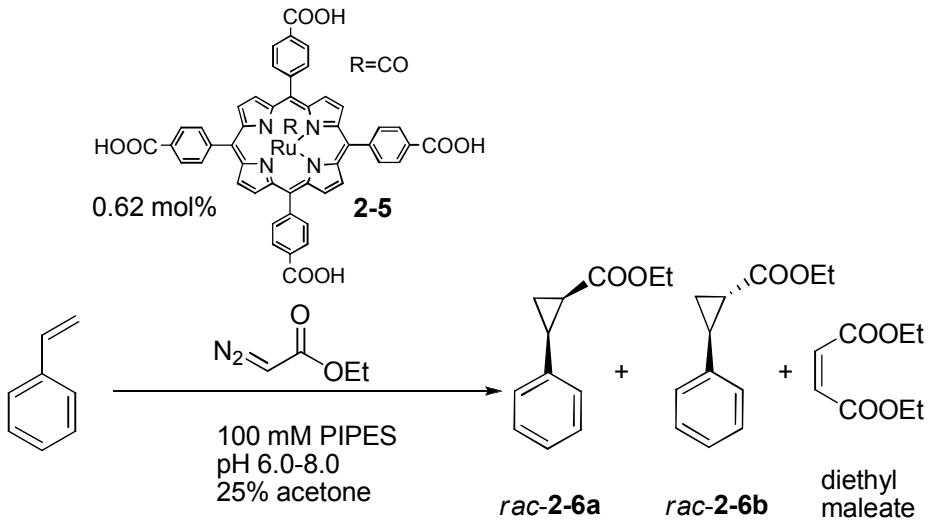
2.3.4 Influence of the pH on the Cyclopropanation Reaction in Aqueous Medium

The influence of the pH on the cyclopropanation was investigated because the stability of the catalytically active species and formation of the desired cyclopropanation products and potential side-products could be pH dependent (see Scheme 2-7). Therefore a series of cyclopropanation reactions of styrene with EDA, catalyzed by ruthenium porphyrin **2-5**, was carried out in buffered solution in the pH range of 6.0 to 8.0. The amounts of the reagent styrene, the side-product diethyl maleate and the desired cis- and trans cyclopropanation products **2-6a,b** and their corresponding yields were determined by GC-FID (Figure 2-5; Table 2-1).

Table 2-1. Yields of the cyclopropanation reaction of styrene with EDA catalyzed by **2-5** and unreacted starting material.

pH	styrene [%]	Yield of diethyl maleate [%]	Yield of **2-6a** [%]	Yield of **2-6b** [%]
6.0	26.31	28.08	0.78	16.22
6.5	17.88	7.19	0.35	27.97
7.0	19.53	7.69	0.38	37.15
7.5	19.18	7.87	0.30	29.15
8.0	18.52	8.25	0.31	29.63

Ruthenium porphyrin **2-5** successfully catalyzed the cyclopropanation of styrene and afforded a fair yield of the trans-product **2-6b** (about 33%) in the pH range of 6.5 to 8.0. At pH 6.0 the catalyst precipitated during the reaction and this presumably led to the observed low yield of about 16%. The reaction did not go to completion because unreacted styrene was found in the reaction mixture. Over the pH range of 6.5 to 8.0 the product distribution remained essentially unchanged despite a 100-fold proton concentration difference (pH 6.0 compared to pH 8.0). Slightly lower yields at acidic pH may be caused by lower local concentrations of styrene due to polymerization. The polymerization of styrene in aqueous medium might even be accelerated by hydrophobic effects (see chapter 1.1.2). At pH 7.5 and 8.0 hydrolysis of the ester moiety of **2-6a,b** could be the reason for slightly lower yields at pH 7.5 and 8.0. The yield of diethyl maleate, requiring the reaction of two molecules of EDA each, is low in the pH range of 6.5 to 8.0.



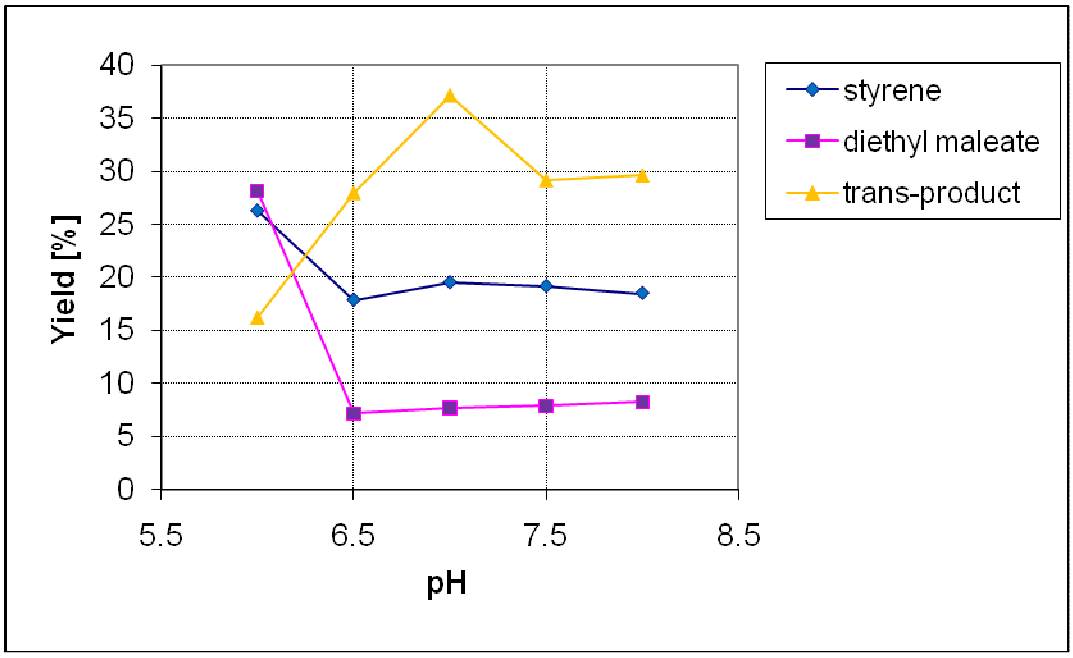


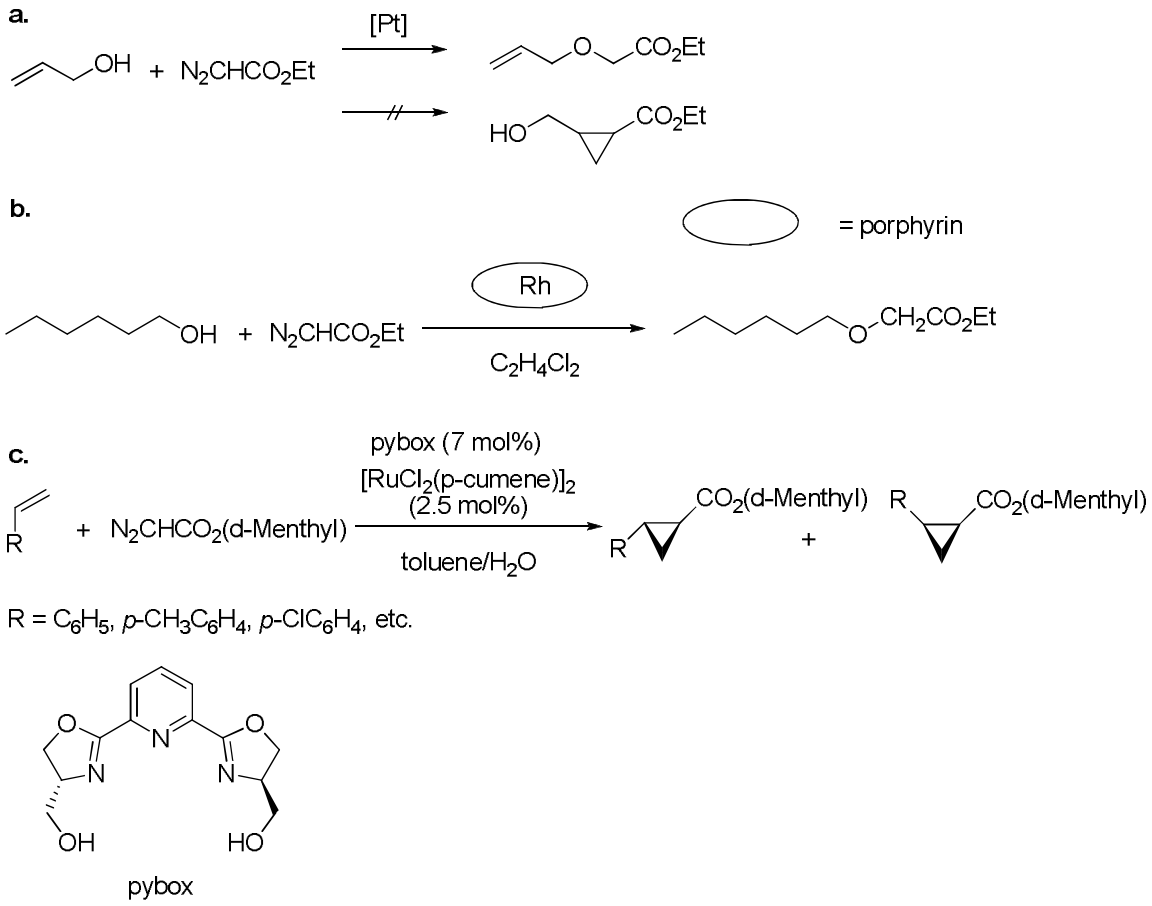
Figure 2-5. Cyclopropanation of styrene in aqueous medium with varying pH.

With regard to the formed *trans*-cyclopropanation product **2-6b** the yield of the maleic ester ranged from 20% to 27%. At pH 6.0 much more diethyl maleate is probably formed due to precipitation of the ruthenium catalyst suggesting that the relative concentration of

catalyst compared to EDA has an impact on the formation of the maleic ester. This is consistent with findings that fast addition of EDA to ruthenium porphyrins yielded increased amounts of maleate.[45] Diethyl maleate is considered to be the kinetic product while **2-6a,b** are considered to be the thermodynamic products.

The successful cyclopropanation with the ruthenium porphyrin catalyst **2-5** in aqueous medium is remarkable. Water as a reaction medium did not deactivate the ruthenium metal center. The proposed catalytically active Fisher-carbene intermediate (Figure 2-1) is either thermodynamically or kinetically stable enough in water to facilitate the cyclopropanation reaction. Decomposition of the carbene intermediate by insertion in the O-H bond of water was not observed; however, other organometallic platinum series metal complexes show different behavior.[46] The addition of ethyl diazoacetate to allylalcohol in the presence of a catalytic amount of platinum complexes afforded the corresponding OH insertion product without yielding any cyclopropanation product (Scheme 2-4a).[47] Rhodium carbene complexes react readily with water by OH insertion. Particularly interesting is the formation of ethers catalyzed by rhodium porphyrins (Scheme 2-4b).[48] Hence, if the ruthenium metal center of catalyst **2-5** was replaced with rhodium then transformation of the diazo moiety of EDA to a hydroxyl group instead of a cyclopropanation reaction would have most likely occurred. However, ruthenium complexes can also show this type of reactivity. Ruthenium complexes with N-(p-toluenesulfonyl)diamine ligands are also efficient catalysts for the insertion of OH bonds into carbenes generated from diazo esters.[49] On the other hand, organometallic ruthenium catalysts can facilitate cyclopropanation reactions in aqueous medium. Due to possible OH insertion it is not surprising that there are only few examples of aqueous cyclopropanation reactions. The first example of a successful cyclopropanation in

aqueous media was achieved by bis(hydroxymethyldihydrooxazolyl)pyridine-ruthenium catalysts (ruthenium-pybox) (Scheme 2-4c).[50] In agreement with the results of the cyclopropanation catalyzed by Ru(CO)TPP-acid in aqueous medium are X-ray structures of ruthenium porphyrin carbenes having hydrophilic solvents like methanol and ethanol as ligands.[26] This suggests that OH insertion is indeed not favorable for carbenes of ruthenium porphyrins. The yields of ruthenium-pybox catalyzed cyclopropanation reactions in water (biphasic) are similar to the obtained yields for the Ru(CO)TPP-acid catalyzed reaction; they usually are around 40 to 50 %. In contrast, Simmoneaux and coworkers catalyzed the cyclopropanation of styrene with EDA by the structurally related Ru(CO)TPP (Figure 2-1) in much higher yields in CH_2Cl_2 (85 %).[26] In summary, the successful cyclopropanation by **2-5** in water with acceptable yields represents an encouraging result for the utilization of ruthenium porphyrins as cofactors in semi-synthetic proteins.



Scheme 2-4.[47,48,50]

2.3.4.1 Mass Balance of Reaction

The sum of the molar equivalents of the products has to equal the sum of the molar equivalents of the starting materials. Summing up of the products and the educts revealed that about one third of the compounds could not be accounted for by GC-FID quantitation (Figure 2-6). The missing compounds are most likely styrene because of losses by polymerization and ethyl diazoacetate due to thermal decomposition in the GC column.

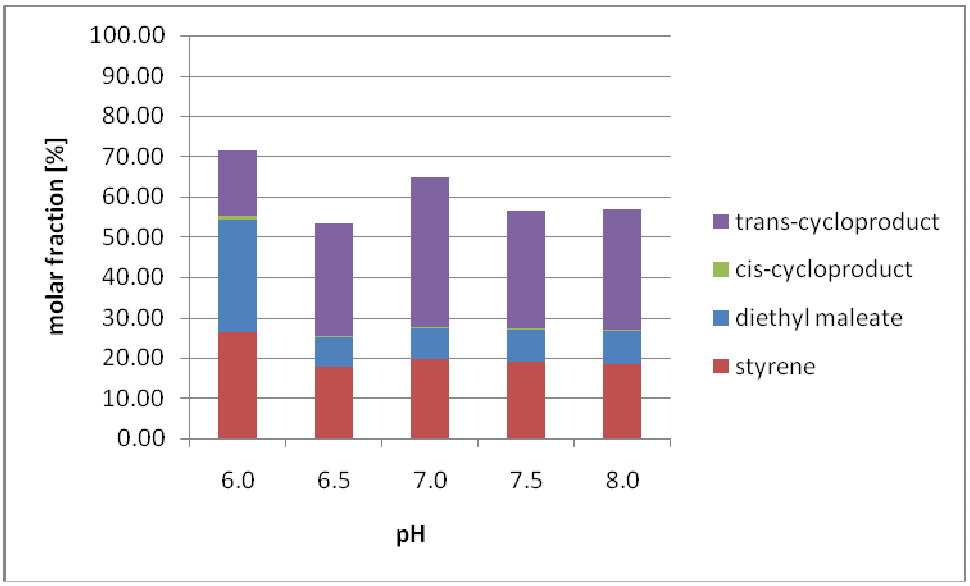


Figure 2-6. Molar fractions of quantitated educts and products.

2.3.4.2 Stereoselectivity of Reaction

The cyclopropanation reactions proceeded with high cis/trans selectivity in the pH range of 6.5 to 8.0 (pH 6.0 was omitted due to precipitation of the catalyst). The cis/trans ratio ranged from 1:80 up to 1:99 (Figure 2-7). This constitutes the highest stereoselectivity reported for the cyclopropanation of styrene with EDA that has been catalyzed by a ruthenium porphyrin. In comparison, Simmoneaux and coworkers catalyzed the same reaction with the structurally similar Ru(CO)TPP (Figure 2-1) in CH_2Cl_2 with a cis/trans ratio of only 1:14 (Figure 2-8).[26] Among organic solvents, such as benzene, toluene and dichloromethane, the product distribution of this kind of reaction varies only slightly. Apparently, the selection of water instead of organic solvents can have a strong impact on the stereoselectivity of the cyclopropanation reaction.

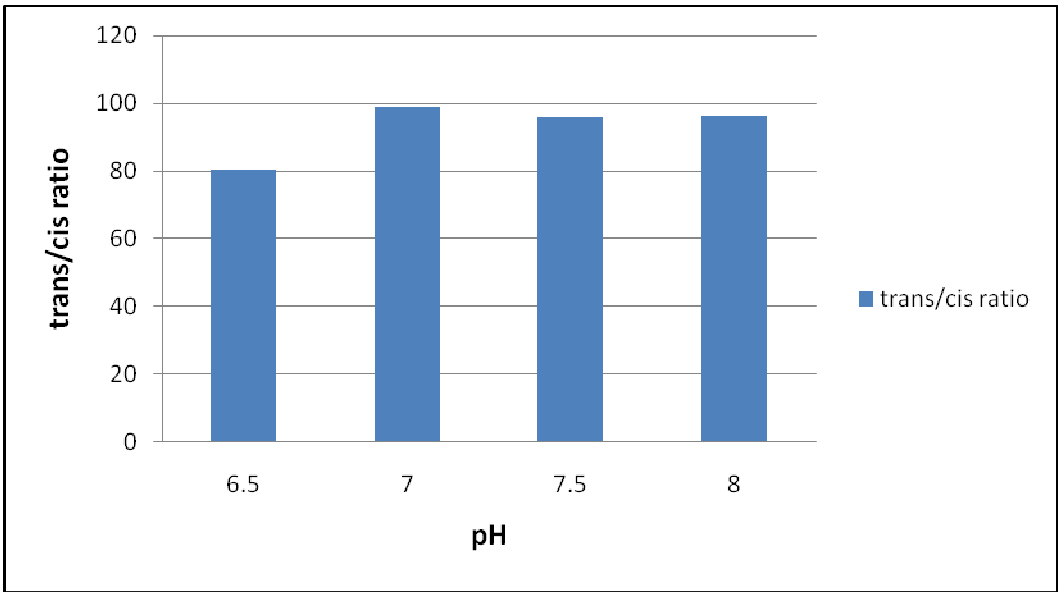


Figure 2-7. Stereoselectivity of cyclopropanation of styrene in aqueous medium with varying pH.

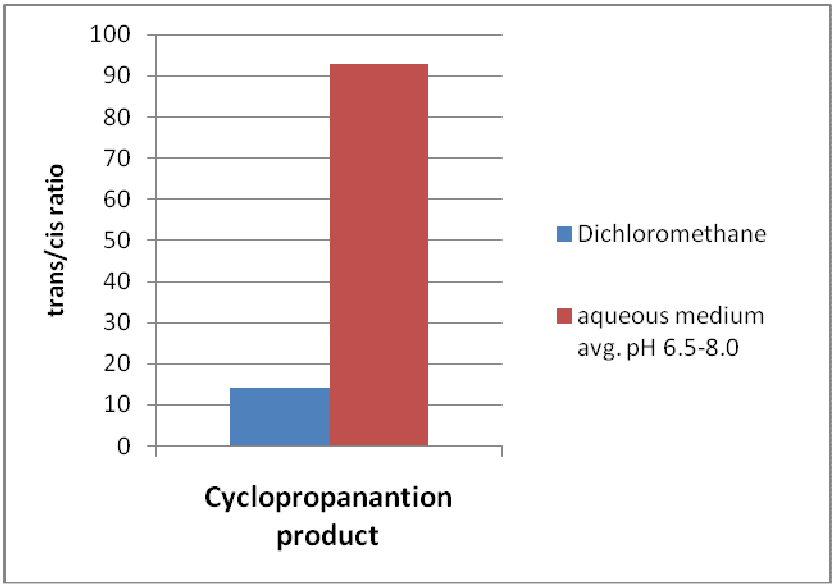
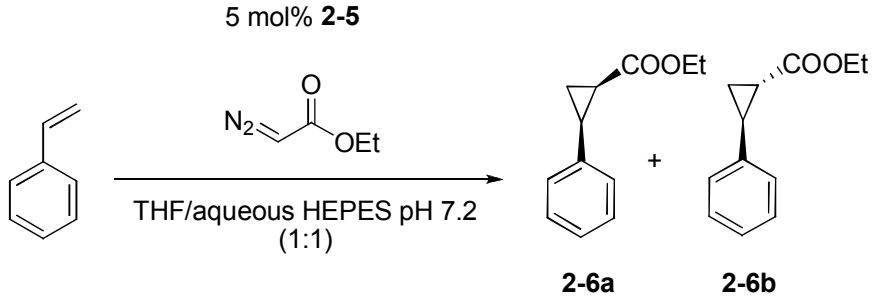


Figure 2-8. Influence of solvents on trans/cis ratios of the cyclopropanation of styrene.

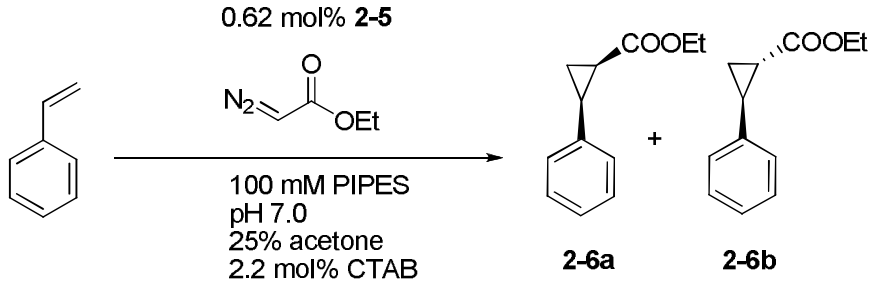
2.3.5 Phase Transfer Catalysis

The catalytic cyclopropanation of styrene was carried out in a biphasic system consisting of an equal mixture of THF and aqueous HEPES buffer.[51] Under these conditions the cyclopropanantion products **2-6a** and **2-6b** were obtained with only 14 % yield (Scheme 2-5). The yield was dramatically improved (83 %) by addition of 5 mol% of the phase transfer catalyst cetyltrimethylammonium bromide.



Scheme 2-5.

Given the increase of yield, styrene was catalytically transformed to its cyclopropanation product with ruthenium porphyrin **2-5** in the presence of 2.2 mol% of the phase transfer catalyst cetyltrimethylammonium bromide (CTAB) (Scheme 2-6). Styrene only dissolves partially in an acetone/water mixture. The yield might increase through better solubility caused by the phase transfer catalyst. Unfortunately, the reaction proceeded in a sluggish way, with the generation of dark, decomposed and insoluble material during the reaction. The reaction yield was only about 9%.



Scheme 2-6.

2.3.6 Reagent Addition Time Studies

The catalytically active species of the cyclopropanation of styrene by ruthenium porphyrin 2-5 is presumably a carbene species, which is formed by reaction with EDA (Scheme 2-7). This carbene species can either react with styrene to yield the desired cyclopropanation product or react with EDA forming diethyl maleate. That is why the formation of diethyl maleate can be reduced or inhibited by keeping the concentration of EDA in the reaction mixture low (compared to styrene) through slow addition.[52] Consequently, several different addition times of ethyl diazoacetate were utilized to catalyze the cyclopropanation of styrene with ruthenium porphyrin **2-5** in order to optimize the reaction conditions (Figure 2-9). The previously arbitrarily chosen EDA addition time of eight hours (used for the pH influence studies) turned out to provide the best yields of cyclopropanation product. Shorter addition times resulted in decreased formation of the desired product and increased diethyl maleate. Longer addition times led to slightly lower yields and about the same amount of the maleic ester side-product (compared to eight hours addition time). Therefore the addition time of eight hours was chosen for all further studies.

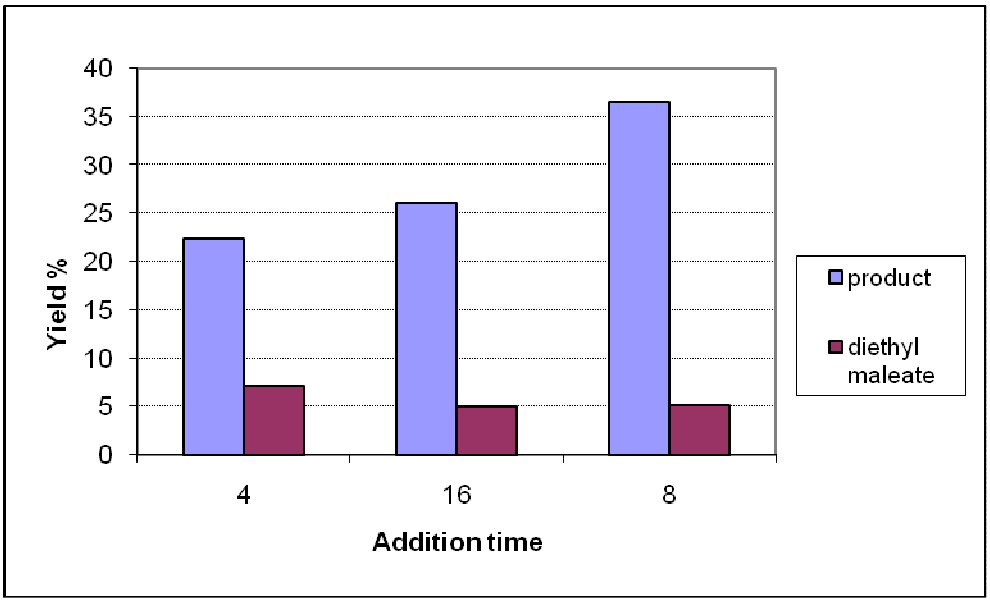


Figure 2-9. Yields of **2-6b** of the cyclopropanation reaction of styrene with EDA catalyzed by **2-5** in comparison with the yields of diethyl maleate. The respective yields were obtained with varying addition times of EDA.

2.3.7 Concentration Dependent Turnover Numbers

A series of cyclopropanations of styrene with EDA in aqueous medium catalyzed by varying amounts of ruthenium porphryin **2-5** were conducted. Catalytic cyclopropanation activity was obtained even at very low catalyst loads (7.6% yield with 0.03 mol% **2-5**) (Figure 2-10). The yield of cyclopropanation product **2-6b** and the amount of utilized catalyst were not directly proportional to each other. Therefore the resulting catalyst load dependent turnover numbers were not constant (Figure 2-11). Lower catalyst loads led to higher turnover numbers. The difference in turnover numbers between the highest and lowest amount of ruthenium porphryin **2-5** was almost five-fold. The catalyst turnovers were in the range of cyclopropanations of styrene with ethyl diazoacetate catalyzed by similar ruthenium carbonyl porphyrins in CH_2Cl_2.[53]

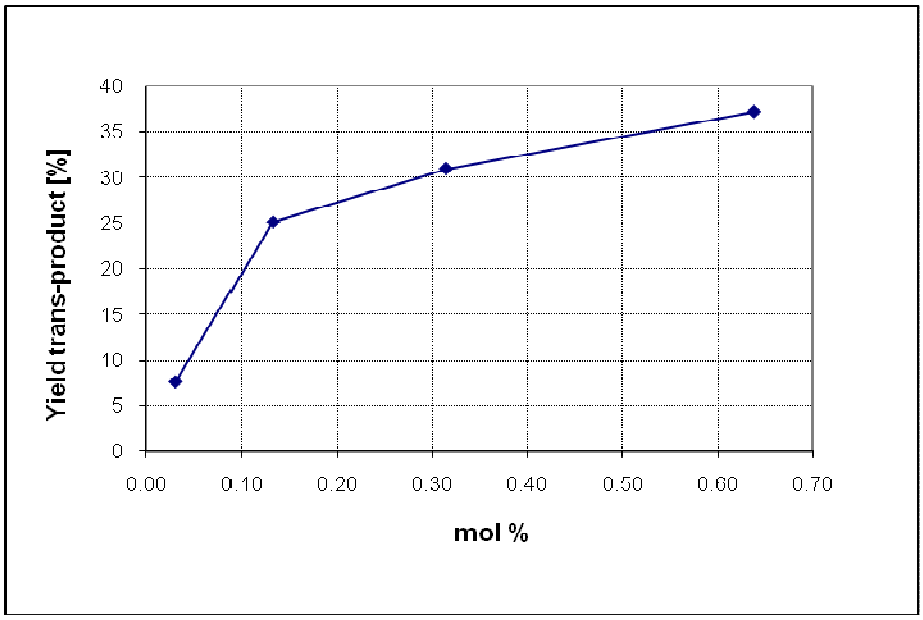


Figure 2-10. Yield dependence of **2-6b** on catalyst load (**2-5**) of the cyclopropanation of styrene.

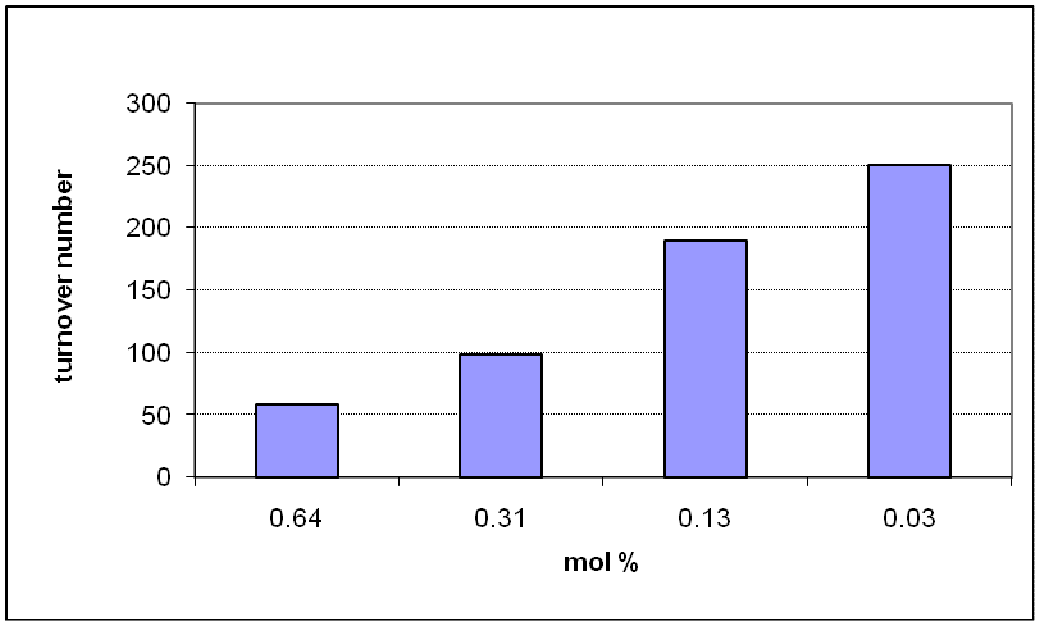
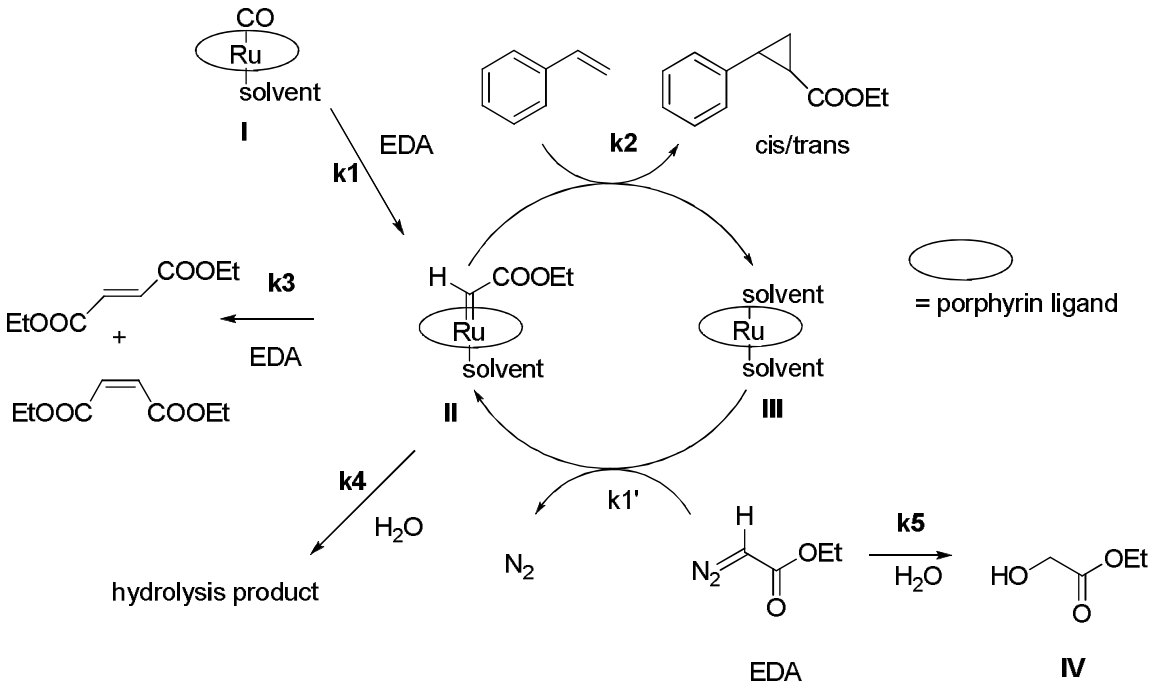


Figure 2-11. Dependence of turnover number on catalyst load (**2-5**) of the cyclopropanation of styrene

2.4 Reaction Kinetics

2.4.1 Catalytic Cycle

In order to analyze potential reaction pathways of the cyclopropanation reaction in water the kinetic parameters of the catalytic pathways had to be determined. A catalytic cycle for the cyclopropanation of styrene with ruthenium porphyrins was designed based on the findings of Che, Collman and Simmoneaux et al. for this type of reaction in organic solvents and potential pathways occurring in aqueous medium added.[23,26,45,53] The catalytic cycle commences with the reaction of the ruthenium carbonyl porphyrin (**I**) with ethyl diazoacetate, forming a carbene species (**II**) while undergoing decarbonylation. The supposedly electrophilic carbene intermediate is attacked by the electron rich olefin styrene yielding the cis/trans cyclopropanation product. In lieu of reacting with styrene carbene species **II** can also undergo reaction with another molecule of EDA producing the olefins diethyl maleate and diethyl fumarate (the formation of diethyl fumarate was never observed in aqueous medium; see chapter 2.3.2). The resulting 6-coordinated ruthenium porphyrin species **III**, having only water as axial ligands, is an active cyclopropanation catalyst. It can react with a new molecule of EDA generating ruthenium porphyrin carbene **II** again, thus concluding the catalytic cycle. Ethyl diazoacetate itself can undergo hydrolysis forming the glycolic acid ester **IV**. This ester might also be formed by OH-insertion into carbene species **II**.



Scheme 2-7.

2.4.2 Kinetics of Carbene Formation

The formation of the carbene-complex for the reaction of **2-5** with ethyl diazoacetate in aqueous solution at neutral pH was monitored by UV-vis spectroscopy (Figure 2-12) (Scheme 2-7). Varying equivalents of EDA (1 to 4) were reacted to examine the kinetics of the reaction (Figure 2-13). Non-linear least squares fit of the UV-vis traces over the entire wavelength range revealed EDA dependent second order rate constants (Table 2-2). The formed carbene species can react with another molecule of EDA yielding diethyl maleate. This reaction pathway causes an error for the observed rate constants for the generation of carbene complex **II**. The more equivalents of EDA are used the more diethyl maleate is formed. Therefore the most accurate rate constant for the formation of **II** is 5.2 ± $M^{-1}s^{-1}$.

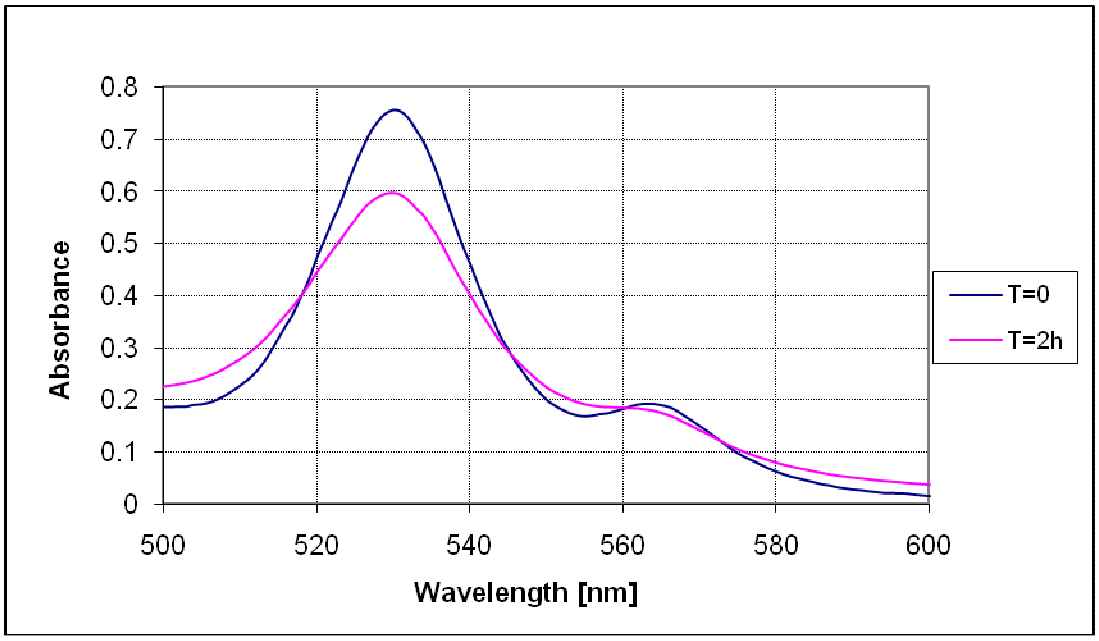


Figure 2-12. UV-vis traces at the beginning (t=0) and after 2 h of the reaction of **2-5** (40 µM) with EDA. The reaction was carried out at room temperature in PIPES buffer at pH 7.0.

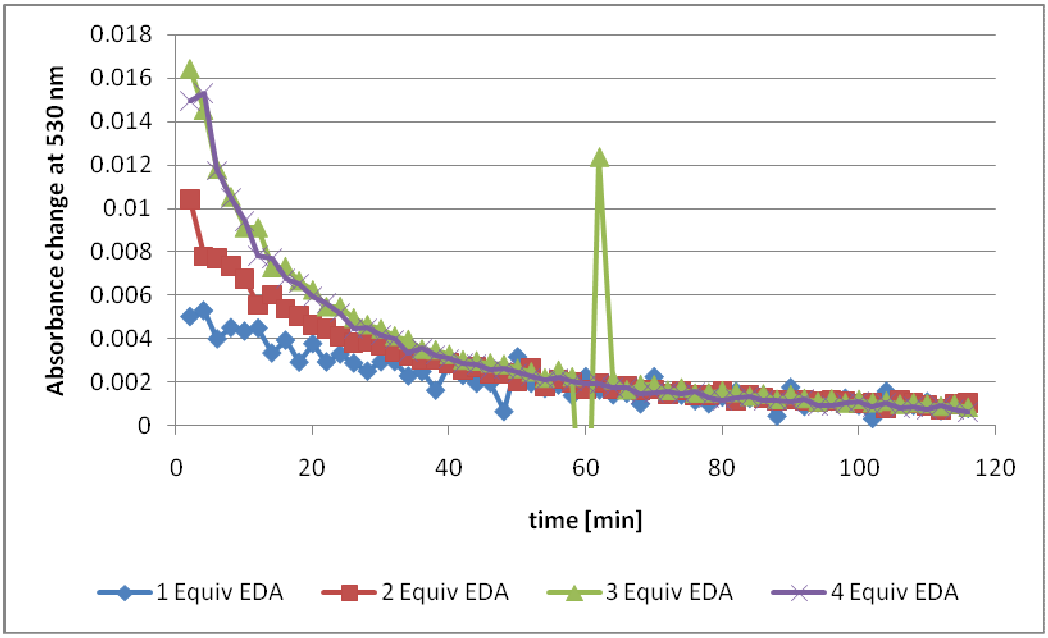


Figure 2-13. Absorbance change at 530 nm as a function of time and added equivalents of EDA. The reaction of **2-5** (40 µM) with EDA was carried out at room temperature in PIPES buffer at pH 7.0.

Table 2-2. EDA dependent rate constants for the aqueous cyclopropanation of styrene.

Equiv. EDA	1	2	3	4
k_2 [$M^{-1}s^{-1}$]	5.2426	4.7682	4.1965	3.4914
Standard deviation	0.07234	0.04584	0.06883	0.071891

$$-\frac{d[EDA]}{dt} = k_2[EDA][catalyst]$$

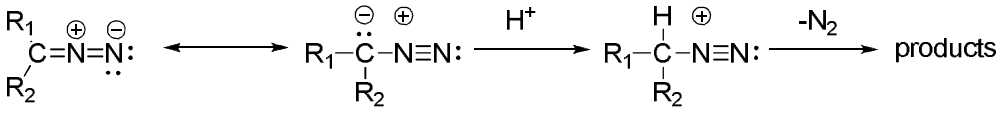
2.4.3 Kinetics of Reaction of Carbene-Complex with Styrene

EDA was reacted with ruthenium porphyrin **2-5** and various equivalents of styrene. Spectral shifts in the visible range could also be observed, but the final absorbance was significantly higher after 2 hours than without styrene. Thus, the rate constant of the reaction of the carbene and styrene was determined by fitting to the following system: A+B⇒C; C+D⇒E+F. However, the data could be fitted to several kinetic systems and different starting values for the rate constant curve fit resulted in different values for k. Therefore no reliable rate constant for the reaction of the carbene with styrene could be obtained. This might be due to the error caused by the poor solubility of styrene in water.

2.4.4 Kinetics of Decomposition of EDA in water

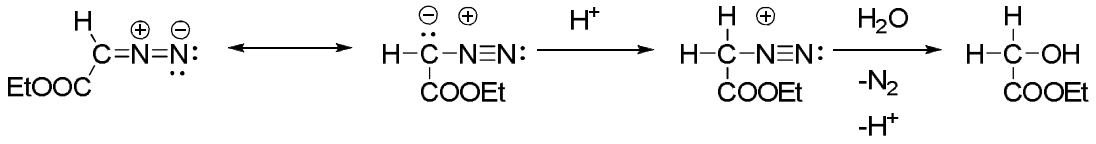
The rate of decomposition of EDA in water has to be determined in order to correct the rate constants for the formation of the carbene-complex and to examine all the kinetic pathways. Instability in the presence of acids is a characteristic property of diazoalkanes.[54] Protonation of the diazocarbon is the crucial first step of diazo

decomposition (Scheme 2-8). After protonation the diazo compound can decompose irreversibly into several different products under release of nitrogen. Staudinger discovered that decreased basicity of diazo compounds leads to decreased rates of decomposition.[55]



Scheme 2-8. Acid catalyzed decomposition of diazo compounds.

Electron-withdrawing groups reduce the nucleophilicity of the diazo carbon (-I effect) and consequently its basicity. Therefore an ester moiety should significantly decrease the rate of decomposition of a diazo alkane. Thus, ethyl diazoacetate is predicted to be fairly stable in water. Roberts el al. studied the decomposition of ethyl diazoacetate in water in the presence of acids (Scheme 2-9).[56] The formation of only the glycolic acid ester **IV** as decomposition product was observed. Ethyl diazoacetate exhibited remarkable stability. The decomposition proceeded quite slowly, even under high concentrations of acids.



Scheme 2-9.

The decomposition of EDA in aqueous solution at neutral pH was monitored by UV-vis spectroscopy. Its strong absorption at 263 nm was used to follow the decomposition (Figure 2-14). After 12 h the change in the absorption spectrum was very small (decrease by ≈ 2 %). Therefore the decomposition of EDA in aqueous medium at neutral pH was considered negligible. However, many compounds can absorb in this range. Therefore, the stability of EDA was also tested by NMR spectroscopy in D_2O and no decomposition was observed after 24 h (data not shown). These results are in agreement with Roberts et al. acid-catalyzed decomposition studies of EDA.[56]

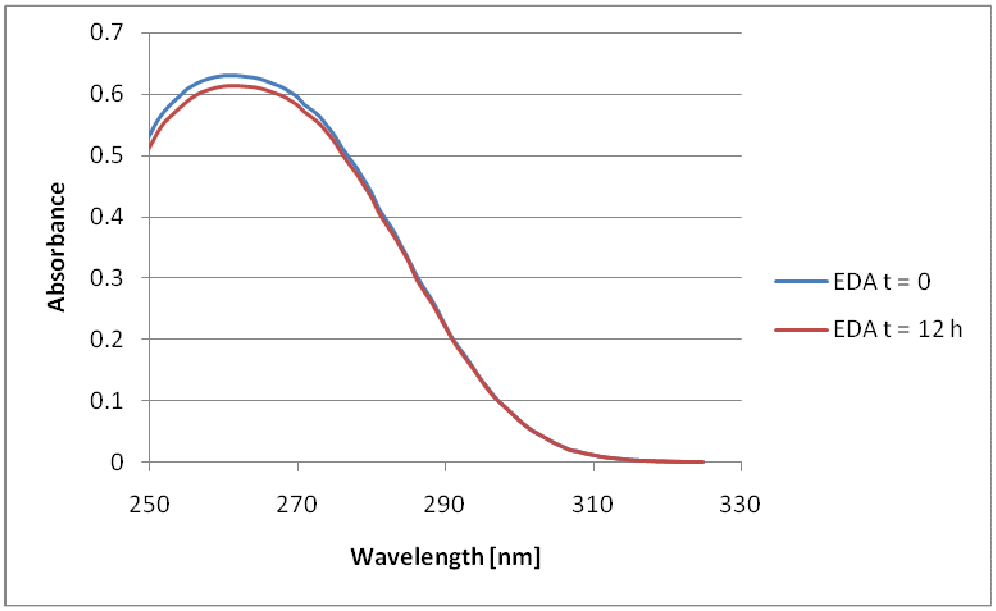


Figure 2-14. UV-vis spectrum of EDA decomposition in water.

2.5 Examination of Reaction Mechanism

2.5.1 Generation of Carbene-Complex in D_2O

It was previously assumed that a carbene-complex intermediate is formed in the reaction of ruthenium porphyrin **2-5** and EDA. In order to confirm this assumption **2-5** was reacted with EDA in phosphate buffered D_2O and the reaction analyzed by NMR spectroscopy (Figure 2-15-A-B). The measured spectra were aligned with the possible side-products ethyl glycolate (Fig. 2-15-C) and diethyl maleate (Fig. 2-16-D) and also with the starting material ethyl diazoacetate (Fig 2-15-E). The carbene-complexes of similar porphyrins, generated by reaction with EDA, were previously described.[26,45,52] The carbene moiety exhibited three signals in C_6D_6: 13.79 (s, C-H, 1H); 2.58 (q, CH_3, 3H); 0.01 (t, CH_2, 3H). The chemical shifts in D_2O and C_6D_6 are usually close to each other unless there are particular interactions between the solvent and the solute, such as hydrogen bonding or π-stacking. However, no peak around 13.8 ppm was found in the spectrum. This could be due to H-D exchange because of the acidity of the proton. The pK_a values of many Fisher carbene complexes were determined by Bernasconi et al.[57] Some of them are rather acidic, e.g. pentacarbonyl(methoxymethylcarbene)chromium(0) was estimated of having a pK_a of 12.3 in water.[58] Unfortunately, the expected triplet and quartet of the carbene moiety (0.01 ppm and 2.58 ppm respectively) were not observed with the correct chemical shift either. The two quartets (q1 and q2) and triplets (t1 and t2) in the obtained spectrum do most likely not derive from a carbene complex because the chemical shifts are off by at least 1.2 ppm from the expected signals at 2.58 ppm and 0.01 ppm. In fact, alignment with several compounds revealed a very good match of the quartet at 4.3 ppm and the triplet at 1.3 ppm with the corresponding peaks of diethyl

maleate. The assignment of the remaining quartet and triplet is not as obvious. The signals match ethyl diazoacetate best but are nevertheless slightly off, thus they cannot be assigned with absolute certainty. Analysis of the reaction mixture by mass spectrometry did not confirm the presence of a carbene complex (data not shown). Due to these findings other reaction and catalyst deactivation pathways have to be taken into account.

In theory, ruthenium complexes can react with diazo compounds in several different ways. The metal center could interact with the lone pair of the terminal nitrogen of the diazo group (Figure 2-16-A). Side-on bonding to the ruthenium metal center is another feasible possibility (Figure 2-16-B). Ruthenium organometallic complexes, which exhibit this type of interactions between the diazo moiety and the metal center, have been isolated and characterized.[59,60] Cenini and coworkers examined the cyclopropanation of styrene with EDA by cobalt porphyrins, which is very similar to the ruthenium porphyrin catalysis.[61] They were also not able to form a terminal catalytically active carbene complex in their NMR experiments. They suggested based on kinetic studies that a porphyrin complex, bound to ethyl diazoacetate with intact diazo moiety, is the actual intermediate that catalyzes the cyclopropanation and not a carbene species. Moreover, it was proposed that diethyl maleate is formed by an N-bridged carbene. It has been shown that iron porphyrin carbene complexes can reversibly isomerize from the terminal carbene form to the N-bridged form and vice versa.[62] In addition, Tatsumi and Hoffmann's theoretical studies of metalloporphyrin carbene complexes show that a decrease of the oxidation state stabilizes the N-bridged form with respect to the terminal one.[63] Besides, ruthenium porphyrins and iron porphyrins are closely related to each

other due to their periodic relationship and can both catalyze the cyclopropanation reaction of styrene with EDA.

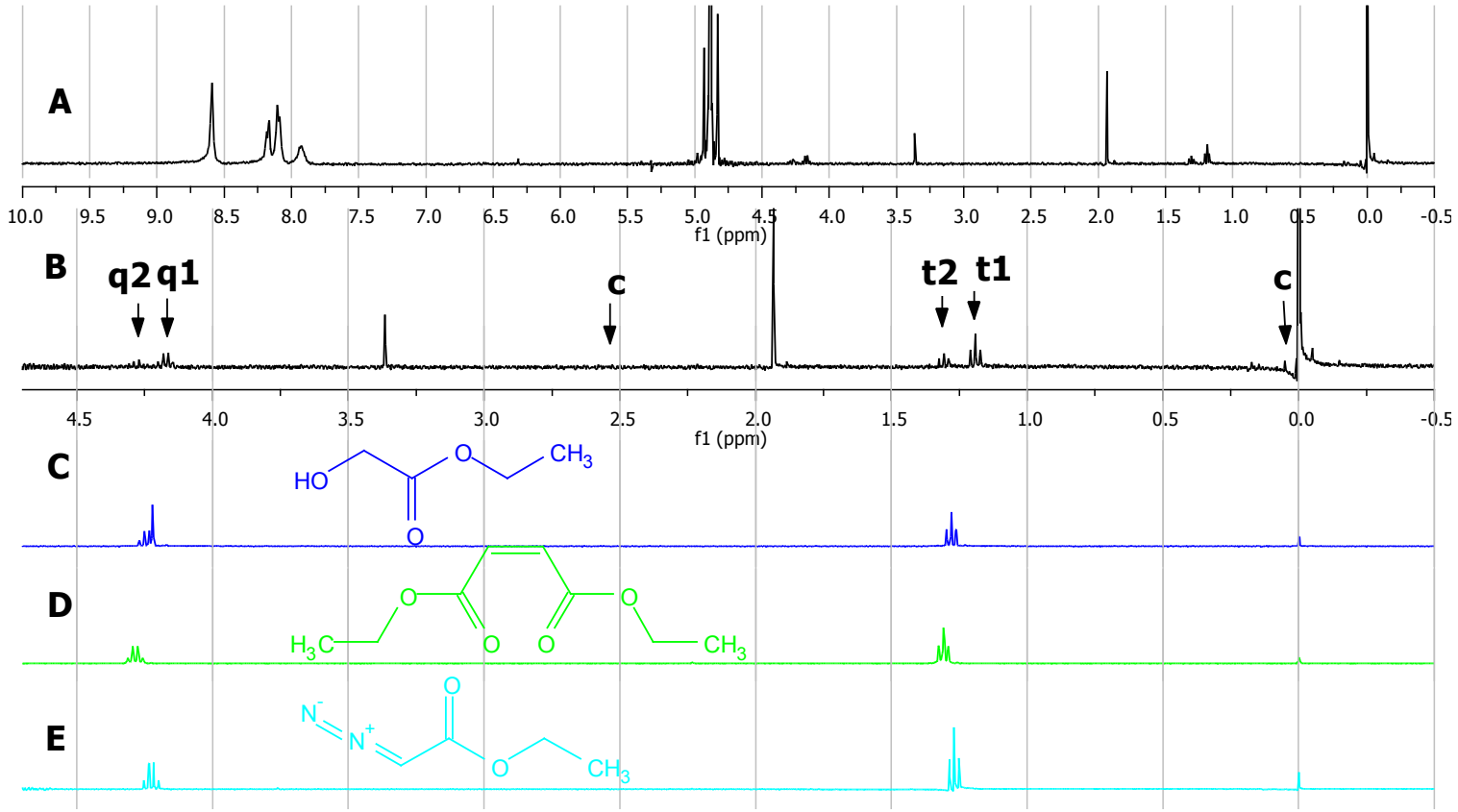


Figure 2-15. ^{1}H-NMR spectra of reaction of EDA with **2-5**. **A** + **B. 2-5** with EDA (1 equiv.). **C.** ethyl glycolate. **D.** diethyl maleate. **E.** EDA. Formed signals by the reaction and the expected signals of the terminal carbene-complex (c) are depicted with arrows.

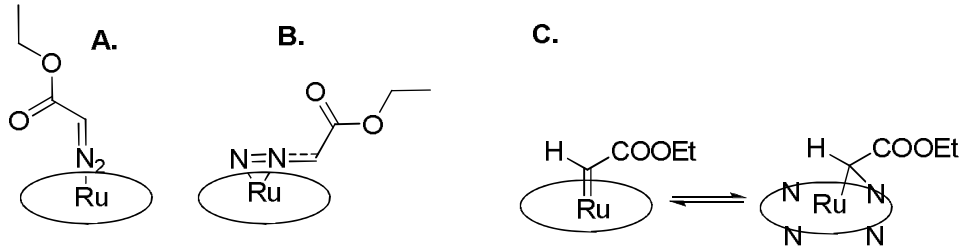


Figure 2-16. Possible complexes formed by the reaction of **2-5** with EDA. **A.** lone pair of terminal nitrogen of the diazo group of EDA coordinating to the ruthenium porphyrin metal center. **B.** side-on coordination of the nitrogens of the diazo group to the ruthenium center. **C.** reversible isomerization of terminal carbene to its N-bridged form.

Hence, ruthenium porphyrins might also be able to form an N-bridged species (Figure 2-16-C); however, the NMR experiments did not confirm the involvement of a bridged carbene form. The ethyl group would be strongly shifted in the spectrum due to the magnetic anisotropy caused by the ring current. The side-on bound complex would show similar behavior and it can therefore be ruled out. If the signals at 1.2 and 3.9 ppm do indeed not belong to EDA then complex **A** (Figure 2-16) would be the most likely match because the ethyl group is probably far enough away from the ring current to exhibit only a small shift in the NMR spectrum. However, all of the discussed complexes might still be present in the reaction but react so fast that they cannot be seen in an NMR spectrum. Hence, the important question if the terminal carbene species is the catalytically active intermediate or simply an “artifact” could not be answered. Moreover, the NMR studies suggest that the measured rate constant for the formation of the carbene complex represents probably the formation of the ruthenium porphyrin species **III** instead of generation of **II** (Scheme 2-7).

2.5.2 Generation of carbene complex in C_6D_6

The terminal carbene complex could not be observed in D_2O in NMR titration experiments. Hence, it was attempted to form the ruthenium carbene species in other solvent systems in order to examine its properties. Collman and coworkers reported the formation of ruthenium carbene complexes with ruthenium porphyrins in C_6D_6.[39,45] That is why ruthenium carbonyl tetraphenyl porphyrin (**2-7**; Figure 2-1), which is structurally very closely related to **2-5**, was reacted with EDA and monitored by NMR spectroscopy Figure 2-17). The carbene complex was successfully formed in C_6D_6 unlike in D_2O (depicted by arrows in Figure 2-17). The signals in the aromatic region changed strongly upon addition of ethyl diazoacetate. Remarkable are the decrease of the signal at 8.9 ppm and the corresponding appearance of a peak at 8.6 ppm. Interestingly, side-products are already formed with one equivalent or less of ethyl diazoacetate. Similar to the reaction in D_2O, comparison of the formed compounds with known side-products and the starting material EDA revealed the formation of diethyl maleate (Figure 2-18). Especially striking are a triplet at 0.75 ppm (with a small hidden triplet at 0.8 ppm) and a quartet at 3.7 ppm (and small quartet at 3.8 ppm probably coupling with the triplet at 0.8 ppm). These signals could not be assigned to any likely compound, such as starting material or known side-products. A similar pattern could be seen in the reaction of **2-5** with EDA in buffered D_2O but it was not clear if the unknown peaks belonged to unreacted ethyl diazoacetate or a new species. Assuming similarities in the reactivities of these ruthenium porphyrins, it can be concluded that the unknown down-shifted triplet and quartet in the reaction in D_2O is not unreacted “free” ethyl diazoacetate (Figure 2-15). The signals might indeed derive from species **A**, **B** or **C** (Figure 2-16).

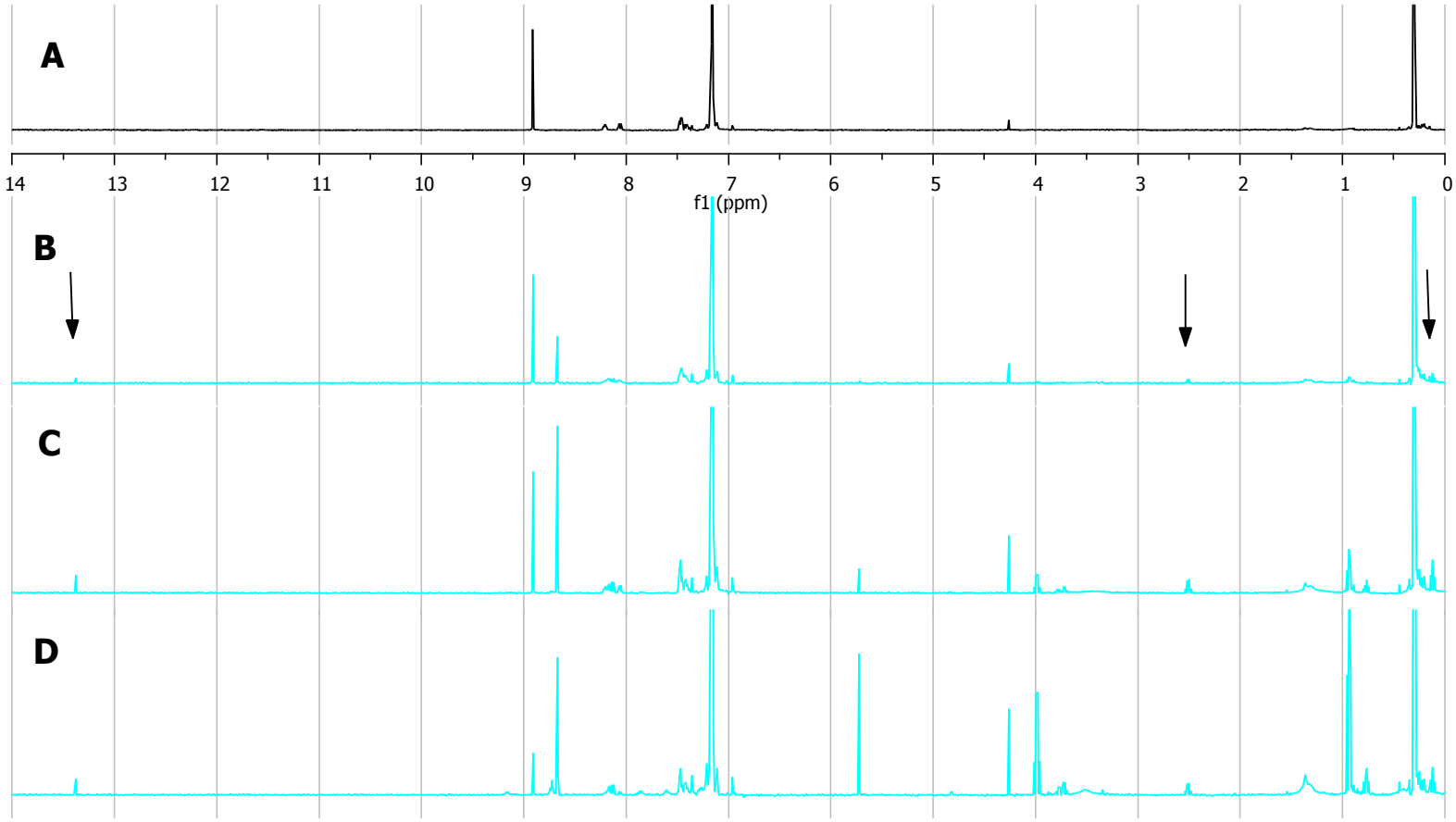


Figure 2-17. ^{1}H-NMR spectrum of carbene formation in C_6D_6. **A.** Ru(CO)TPP (**2-7**) **B. 2-7** + EDA (0.5 equiv.) **C. 2-7** + EDA (1 equiv.). **D. 2-7** + EDA (3 equiv.). The formed signals of the carbene complex are depicted with arrows.

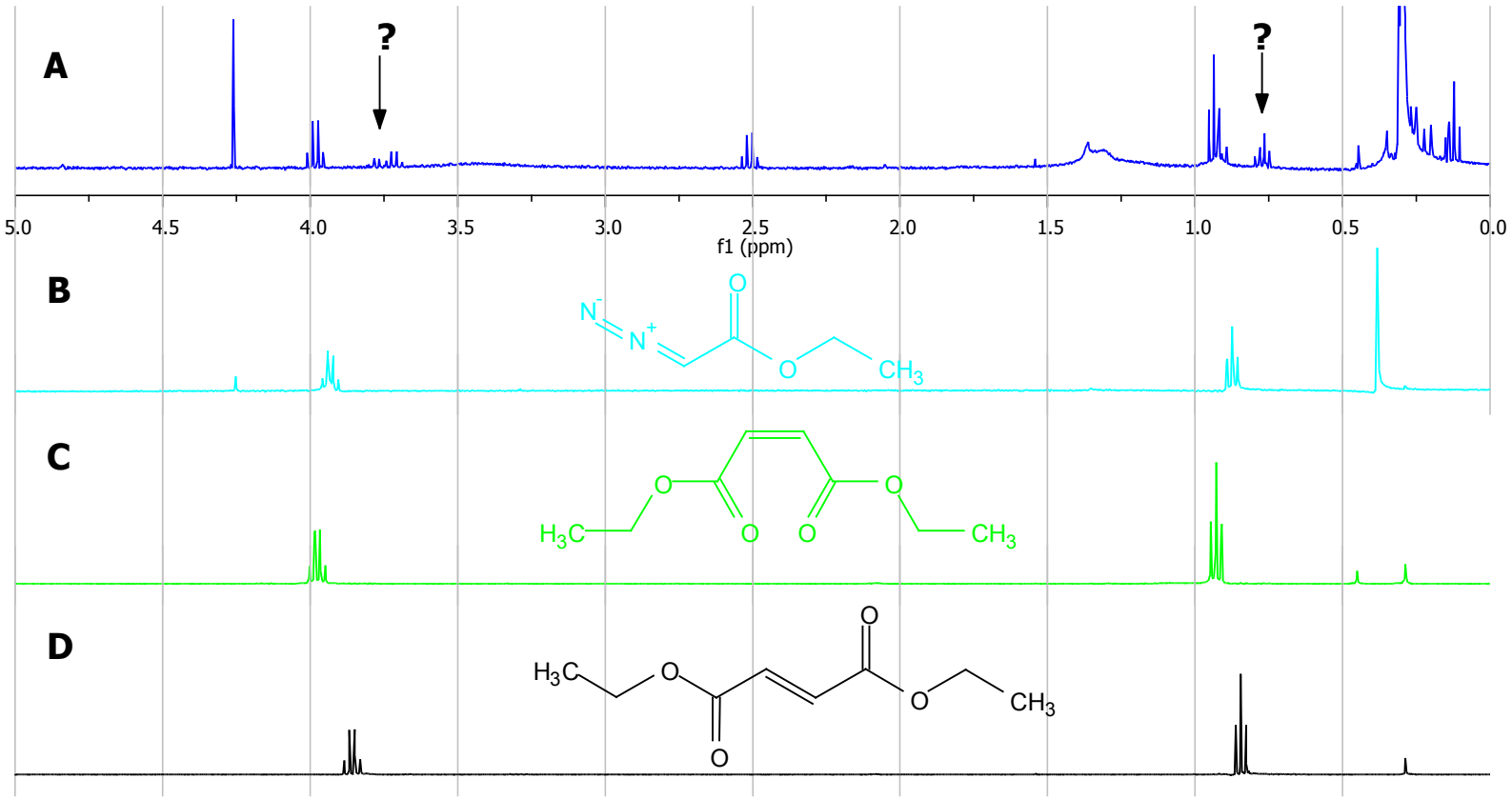


Figure 2-18. ^{1}H-NMR spectrum of carbene formation in C_6D_6. **A. 2-7** + EDA (1 equiv) **B.** EDA **C.** diethyl maleate **D.** diethyl fumarate. Signals that could not be assigned to any likely compound, such as starting material or known side-products, were depicted with arrows.

2.5.3 Influence of Axial Ligands on Carbene Formation

The influence of axial ligands on the formation of the ruthenium porphyrin carbene complex was examined in order to attain a better understanding of important reaction parameters for the cyclopropanation with a semi-synthetic protein. Lewis bases such as imidazole, pyridine, N,N-dimethylamino-pyridine (DMAP) and mercaptoethanol were chosen as axial ligands because they show good resemblance to the axial ligand moieties of heme groups in heme proteins (histidine and cysteine). Thus, they are important models for the feasibility of cyclopropanation reactions by semi-synthetic proteins. The studies were carried out in the non-coordinating solvent C_6D_6 with ruthenium porphyrin **2-7** and monitored by ^{1}H-NMR spectroscopy. The selected axial ligands are stronger π-donors than water and are therefore expected to bind tighter to the ruthenium metal center. Consequently, the opposing bond(s) of either the carbonyl or the carbene moiety should get more weakened (trans-effect). Unfortunately, imidazole could not be successfully used for the axial ligand experiments due to its surprisingly poor solubility in benzene. Ruthenium porphyrin **2-7** was reacted with several equivalents of pyridine (4 equivalents total) (Figure 2-19). Upon binding to the ruthenium metal center pyridine was expected to exhibit three new signals due to the symmetry of the formed complex; however, more than three new peaks appeared in the spectrum. Binding of the axial ligand can also be seen by the slight shift of the porphyrin signal at 8.8 ppm. Interestingly, addition of ethyl diazoacetate did not result in the generation of the characteristic peaks of the carbene complex. Instead, comparison with likely compounds showed a perfect match for the starting material EDA. Apparently, pyridine as axial ligand completely inhibits the reaction of the ruthenium catalyst with ethyl diazoacetate. Reacting **2-7** with DMAP revealed a similar pattern (Figure 2-20). Addition

of DMAP yielded more than the expected amount of new signals and a slightly messy spectrum in the aliphatic region. Again, the carbene formation reaction was inhibited by the axial ligand because only unreacted EDA was found in the spectrum. Mercaptoethanol is a weaker binding ligand and led to completely different results (Figure 2-21). Addition of mercaptoethanol to the ruthenium porphyrin solution in benzene resulted in a fairly messy spectrum in the aliphatic region. Despite the broad peaks due to mercaptoethanol it could be observed that the addition of ethyl diazoacetate did not simply yield in an inert catalyst. Comparison with probable molecules demonstrated that EDA was converted to diethyl maleate. However, no signals of the terminal carbene species were found in the NMR spectrum. Interestingly, traces of the previously observed triplet and quartet (0.75 and 3.7 ppm respectively) of an unknown species were also observed. In summary, the results of the influence of the axial ligand on the formation of the carbene species present a possible obstacle for the development of a semi-synthetic “cyclopropanase” because axial ligands are usually bound to the porphyrin metal center in heme proteins. The limitation of this model is that an excess of ligand is required to ensure that ligand is bound to the metal center. This leads to the potential formation of complexes with a 2:1 binding stochiometry, which are not representative for the ruthenium porphyrin, embedded in the active site of a heme protein.

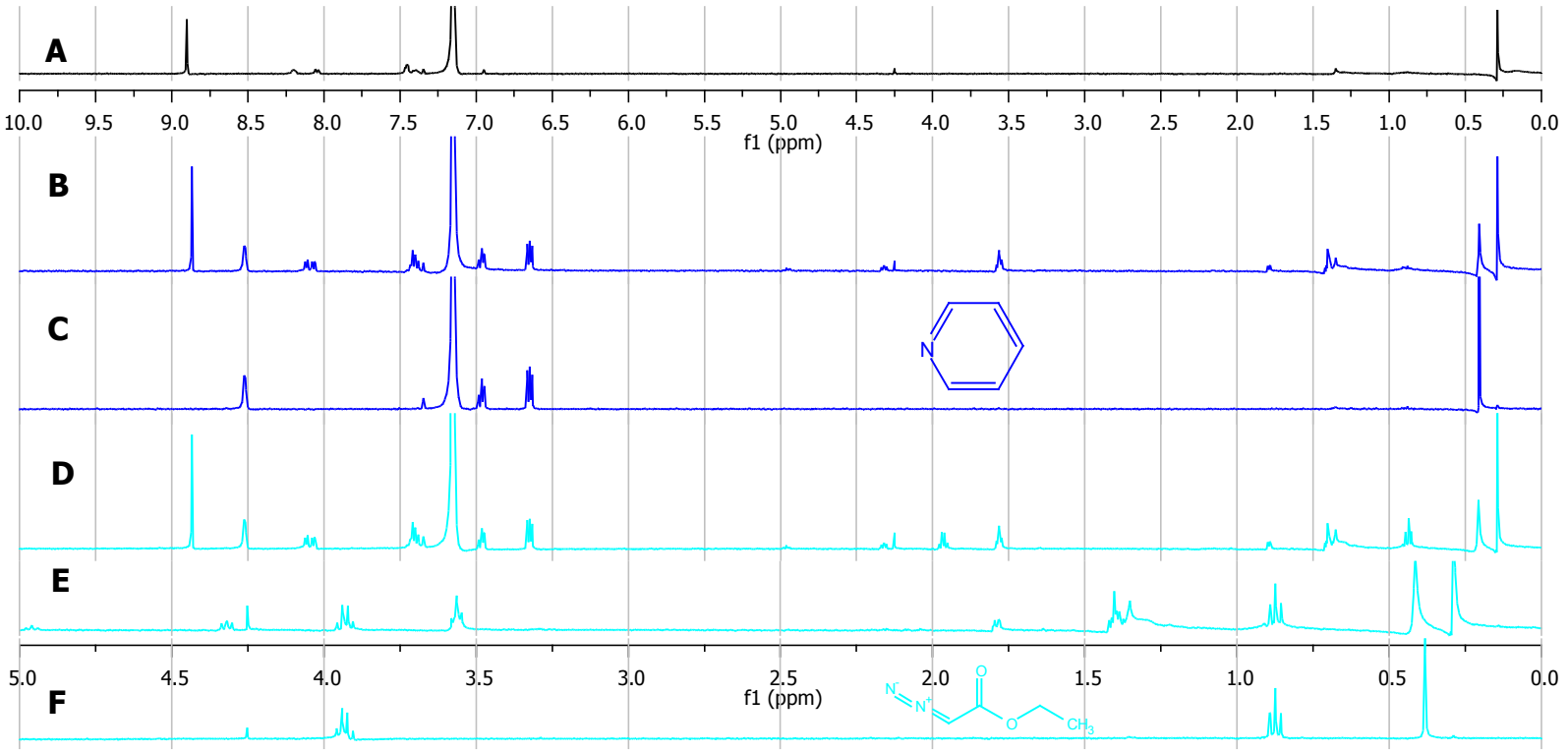


Figure 2-19. ^{1}H-NMR spectrum of carbene formation with pyridine as axial ligand. **A.** Ru(CO)TPP (**2-7**) **B.** 2-7 + pyridine (4 equiv.) **C.** pyridine **D + E. 2-7** + pyridine + EDA (1 equiv.) **F.** EDA

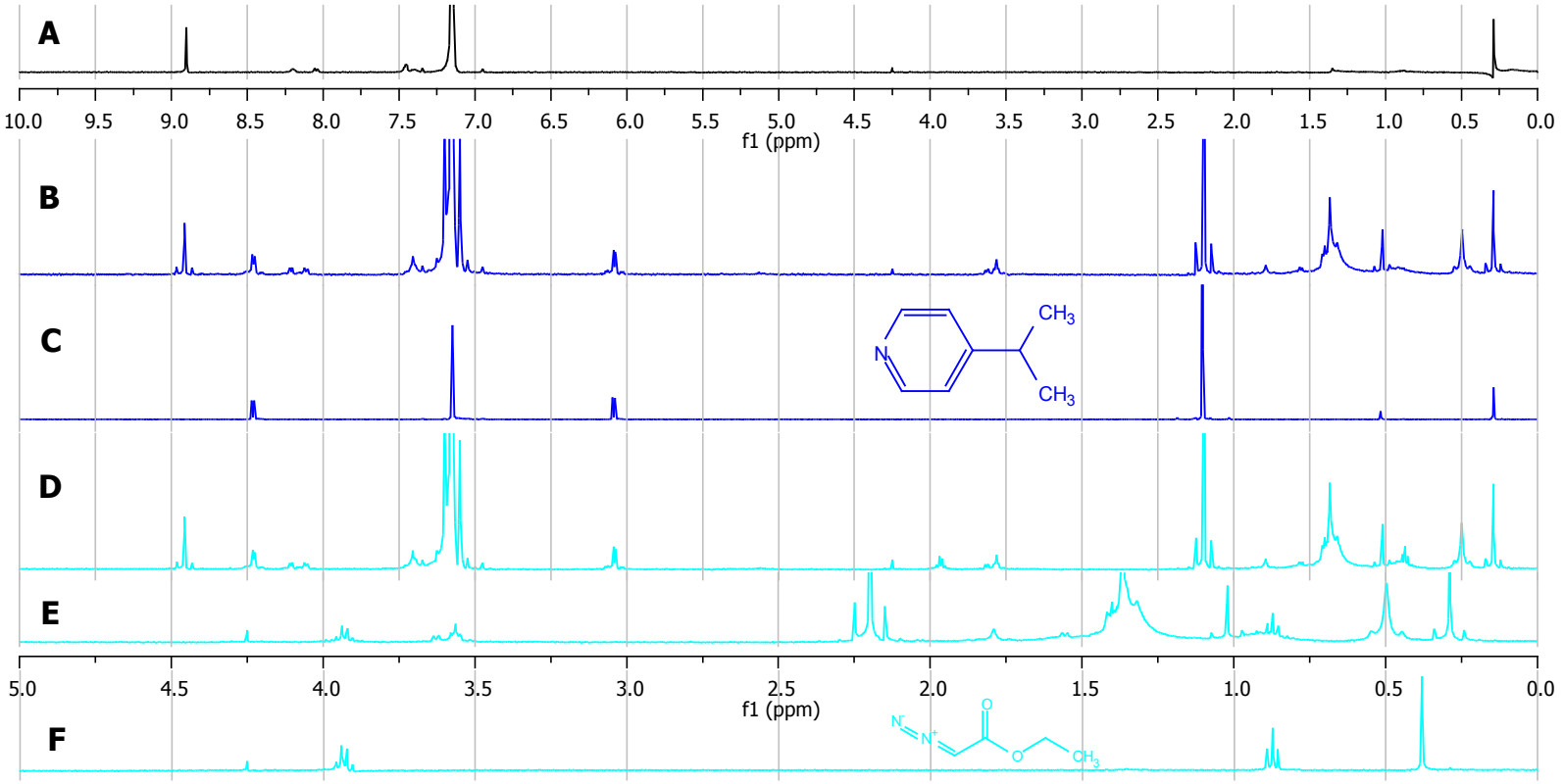


Figure 2-20. ^{1}H-NMR spectrum of carbene formation with DMAP as axial ligand. **A.** Ru(CO)TPP (**2-7**) **B.** 2-7 + DMAP (4 equiv.) **C.** DMAP **D + E. 2-7** + DMAP + EDA (1 equiv.) **F.** EDA

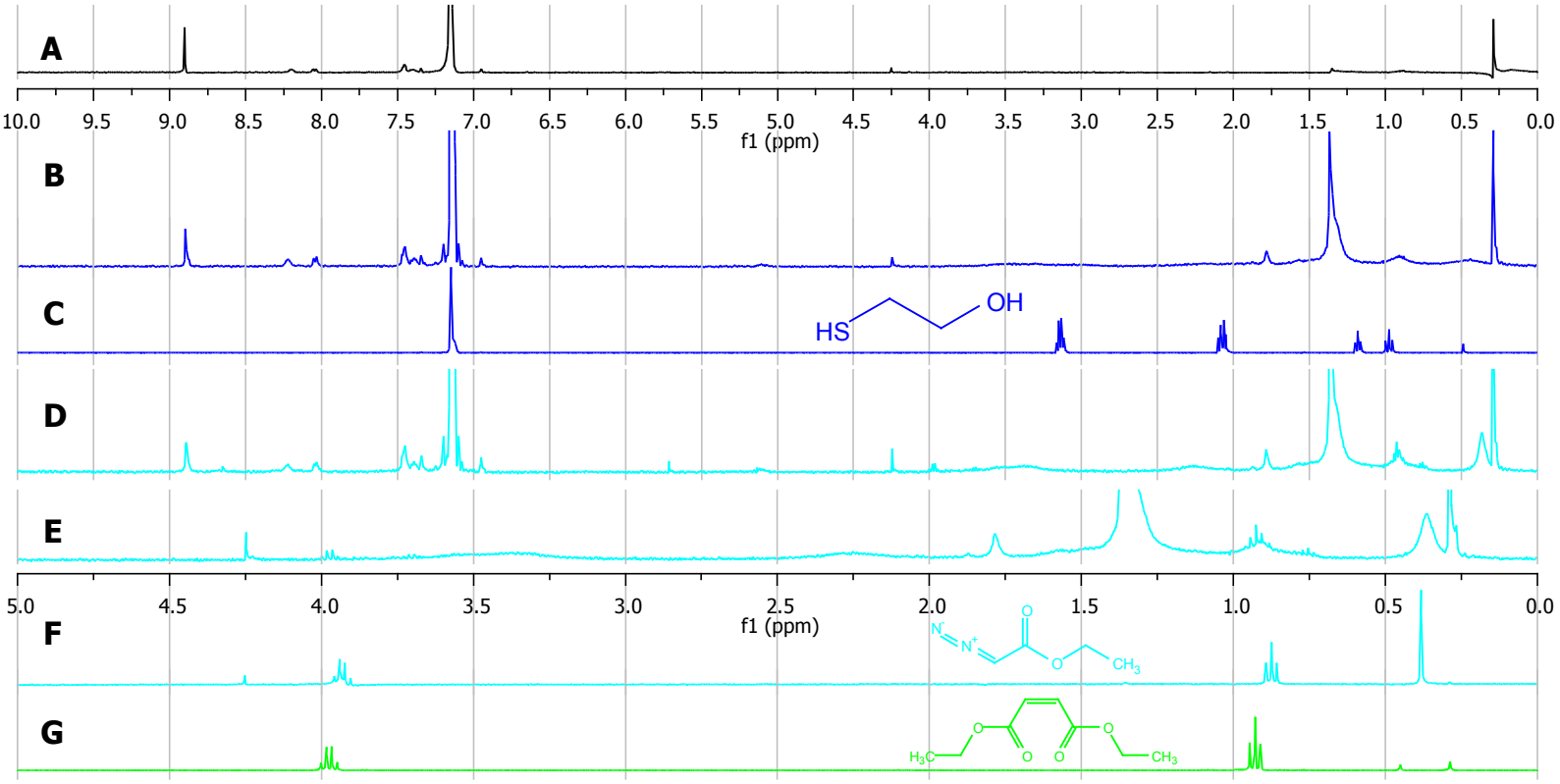


Figure 2-21. [1]H-NMR spectrum of carbene formation with mercaptoethanol as axial ligand. **A.** Ru(CO)TPP (**2-7**) **B.** 2-7 + mercaptoethanol (3 equiv.) **C.** mercaptoethanol **D + E. 2-7** + mercaptoethanol + EDA (1 equiv.) **F.** EDA **G.** diethyl maleate

2.5.4 Examination of the Reaction of Lewis Bases with Ruthenium Carbonyl Porphyrins

It has been shown that the reaction with an excess of certain lewis bases and ruthenium porphyrin **2-7** in C_6D_6 resulted in the emergence of more new signals than expected (Figure 2-19). Therefore the reaction of pyridine with **2-7** was further investigated by electrospray mass spectrometry (MS-ESI), which is a mild ionization method that usually causes only little fragmentation. Ruthenium porphyrin **2-7** was reacted with an excess of pyridine (20 equivalents) at 50 ℃ for several hours (Figure 2-22). Three different cone voltages were used to generate three different levels of ionization (low = 5V; medium = 20V; high = 60V). In all conditions the expected signal of a 1:1 pyridine-ruthenium carbonyl porphyrin complex (**A**) could be observed. However, another complex was also already formed under mild ionization conditions. The exhibited pattern matches the isotope distribution pattern of a 2:1 pyridine-porphyrin complex (**C**) well. The suggested structure of this unknown complex is depicted as a complex with one pyridine that is directly bound to the metal center and a pyridine moiety that is bound to the carbonyl carbon forming an amide-type bond. This complex was stable enough to be apparent at a high cone voltage. The actual ratio of formed **A** to **C** might be quiet different than illustrated in the mass spectrum because the proposed complex **C** is already charged and thus will already exhibit a signal without any additional ionization. It is possible that complex **C** was formed by the MS process and not during the previous reaction. However, the appearance of the signal for complex **C** already at a low cone voltage renders this assumption unlikely. It is noteworthy that the spectrum shows almost no signs of decarbonylation. Hence, pyridine does not bind strongly enough to the ruthenium metal center to induce the loss of the carbonyl group.

The strong signal at high cone voltage can be assigned to an acetonitrile-ruthenium carbonyl porphyrin species (**B**). The acetonitrile derived from the standard ESI mixture utilized (acetonitrile/formic acid) by the Georgia Tech Mass Spectrometry Facility. Only under fairly harsh conditions was the large excess of acetonitrile able to replace some of the axial ligands of **2-7**. Complex **A** exhibited a much greater abundance with regard to complex **C** at high cone voltage. It is not clear if this originates from increased ionization of complex **A**, decomposition of complex **C** or a mixture of both events.

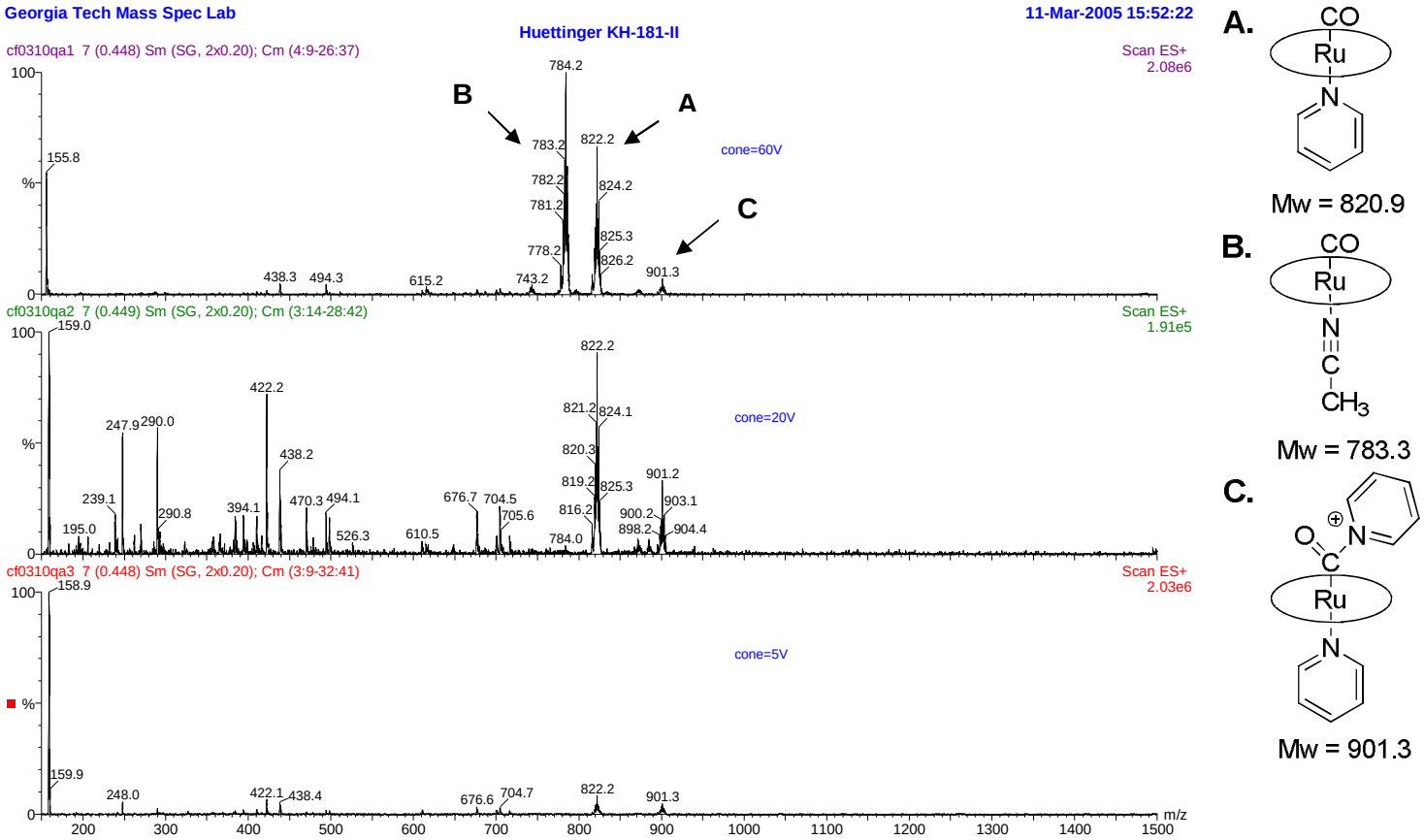


Figure 2-22. MS spectrum of **2-7** reacted with excess of pyridine. **A**,**B** and **C** represent the structures of the most abundant signals.

2.6 Conclusion

This chapter represents the design, development and characterization of a suitable reaction, reaction conditions and catalytic system for later utilization in a semi-synthetic protein. The chosen cyclopropanation reaction proceeded with fair yields and high stereoselectivity in the required aqueous medium. The successful cyclopropanation in aqueous medium constitutes a crucial requirement for a catalytically active semi-synthetic protein. Mechanistically, it remains unclear what the actual catalytic species for the formation of diethyl maleate and the cyclopropanation product are; however, new insights were gained from the mechanistic studies. Moreover, the examination of the influence of the axial ligand on the carbene formation illustrated that Lewis bases, such as pyridine and DMAP, can inhibit the catalytic reactivity. This might be a major obstacle for the development of a catalytically active cyclopropanase because histidine or cysteine is usually coordinated to the metal center of the heme-porphyrin.

2.7 Experimental Section

2.7.1 Materials and Reagents

Methyl 4-formyl benzoate (Aldrich, 99%), pyrrole (Aldrich, 98%), anhydrous decaline (Aldrich, 99% mixture cis + trans), styrene (Aldrich, 99%), ethyl diazoacetate (Aldrich, contains 10% CH_2Cl_2), diethyl maleate (Aldrich, 97%), Ru_3CO_{12} (Acros, 98%). All chemicals, except pyrrole, were used without further purification. NMR: δ in ppm versus SiMe4 (0 ppm, ^{1}H, 400 MHz). MS: selected peaks; m/z. Flash chromatography: Merck silica gel (240-400 mesh). TLC: 0.25 mm, Merck silica gel 60 F_{254}, visualizing at 254 nm or with permanganate solution (2 % $KMnO_4$, 5 % K_2CO_3).

2.7.2 Synthesis

Meso-tetrakis-(4-carboxyphenyl)porphyrin tetramethyl ester (**2-3**) was synthesized following the published procedure.[40]

Ruthenium(II)carbonyl *meso*-tetrakis-(4-carboxyphenyl)porphyrin tetramethyl ester (2-4). Modifications of the procedure of Rillema et al. were used:[41] A mixture of **2-3** (536 mg, 0.49 mmol) and Ru_3CO_{12} (307 mg, 0.48 mmol) in 115 mL of decaline were heated for 5 hours under reflux and then allowed to cool to room temperature. The suspension was filtered and the remaining solid extracted several times with a 100 mL CH_2Cl_2-acetone mixture (96:4). The combined extracts were concentrated under reduced pressure and the crude product was purified by using flash chromatography on silica gel (CH_2Cl_2/acetone 96:4, gradually changed to CH_2Cl_2/acetone 9:1) providing **2-4** as an orange powder (351 mg, 74%). ^{1}H NMR (DMSO-d_6, 400 MHz): δ 8.57 (s, 8H);

8.39-8.34 (m, 12H); 8.25-8.23 (m, 4H); 4.08 (s, 12H). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.66 (s, 8H); 8.41-8.36 (m, 12H); 8.28-8.24 (m, 4H); 4.07 (s, 12H).

Ruthenium(II)carbonyl *meso*-tetrakis-(4-carboxyphenyl)porphyrin (2-5). Modifications of the procedure of Nimri et al. were used:[42] Ruthenium porphyrin **2-4** (200 mg, 0.31 mmol) was added to a 60 mL mixture of 3 M aqueous KOH and methanol (1:1) and refluxed for 4 h. The solution was acidified with acetic acid and extracted several times with EtOAc until the remaining solution was almost colorless. The combined organic extracts were dried over $MgSO_4$ and concentrated under reduced pressure to provide **2-5** as a dark red powder (177 mg, 95% yield). ^{1}H NMR (DMSO-d_6, 400 MHz): δ 8.58 (s, 8H); 8.37-8.32 (m, 12H); 8.22-8.20 (m, 4H).

***trans*-1-carbethoxy-2-phenylcyclopropane** (**2-6b**). Modifications of the procedure of Doyle et al. were used.[64] Ethyl diazoacetate (2.70 mL, 26.0 mmol) was added at a controlled rate over 4 h via syringe pump to a stirred mixture of $Pd(PPh_3)_4$ (100 mg, and styrene (43.72 g, 420 mmol). The reaction mixture was filtered through a pad of celite and concentrated under reduced pressure. The crude product was purified via preparative thin layer chromatography (Chromatotron) on silica gel plates (hexanes/acetone 40:1) to provide a white solid. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.33-7.08 (m, 5H); 4.18 (q, J = 7.2 Hz, 2H); 2.52 (m, 1H); 1.90 (m, 1H); 1.60 (m, 1H); 1.31 (m, 1H); 1.28 (t, J = 7.2 Hz, 3H).

2.7.3 GC-FID and GC-MS

An HP 5890 GC equipped with a 6890 series autosampler was used for GC-FID analysis with an HP-5 capillary column ((5%-phenyl)-methylpolysiloxane). An Agilent GC Chemstation was utilized for peak integration and data analysis. Response factors were calculated for all of the major compounds according to chapter 2.7.5. GC-MS analysis was conducted on an HP 5973 MSD coupled with an HP 5890 GC, using an HP-5 capillary column. Peak identifications were made using the Wiley Registry of Mass Spectral Data (6^{th} edition).

2.7.4 Model Cyclopropanation Reaction for Analysis by GC

Ethyl diazoacetate (90 µL, 0.86 mmol), dissolved in 500 µL acetone, was added at a controlled rate over 12 h via syringe pump to a stirred solution of ruthenium porphyrin **2-5** (5 mg, 5.44 µmol, 0.63 mol%) and styrene (100 µL, 0.87 mmol) in 9 mL PIPES buffer at ambient temperature. The solution was extracted twice with 5 mL diethylether. The GC-MS samples were taken by diluting with acetone and addition of dodecane.

2.7.5 Standardization of Analytes with Internal Standard by GC-FID and Quantitation

A series of solutions with varying amounts of styrene, diethyl maleate and 2-6b in acetone was prepared in the concentration range of 0.1 mg/mL – 1.0 mg/mL. The internal standard dodecane was kept at a concentration of 0.5 mg/mL. The response

factor was determined as the slope of the linear regression line of the mass analyte (x) to mass internal standard ratio (dod = dodecane) against peak area analyte to peak area internal standard ratio (equation 1). The yield could then be calculated according to equation 2 (Mw = molecular weight; n = chemical amount).

$$\frac{mass(x)}{mass(dod.)} = \frac{area(x)}{area(dod)} \times Rf \qquad \text{(equation 1)}$$

$$Yield\% = \frac{\frac{area(x)}{area(dod.)} \times Rf \times mass(dod.)}{Mw(x) \times n(EDA)} \times 100 \qquad \text{(equation 2)}$$

2.7.6 Influence of pH on Cyclopropanation Reactions

Ethyl diazoacetate (90 µL, 0.86 mmol), dissolved in 500 µL acetone, was added at a controlled rate over 8 h via syringe pump to a stirred solution of ruthenium porphyrin **2-5** (5 mg, 5.44 µmol, 0.63 mol%) and styrene (100 µL, 0.87 mmol) in 9 mL PIPES buffer (pH 6.0, 6.5, 7.0, 7.5, 8.0) and 3 mL acetone at ambient temperature. The solution was extracted five times with 5 mL diethylether. The GC-FID samples were taken by diluting with acetone and addition of n-dodecane.

2.7.7 Phase Transfer Catalysis

The procedure of 2.7.6 was followed. The pH 7.0 was chosen and CTAB (7.0 mg, 19 µmol, 2.2 mol%) added to reaction mixture before addition of ethyl diazoacetate.

2.7.8 Reagent Addition Time Studies

The procedure of 2.7.6 was followed at pH 7.0 with addition times of 4 h, 8 h and 16 h of EDA.

2.7.9 Concentration Dependent Turnover Numbers

The procedure of 2.7.6 was followed at pH 7.0 with catalyst loads (**2-5**) of 0.03, 0.13, 0.31 and 0.62 mol%.

2.7.10 Kinetics of Carbene Formation

Ruthenium porphyrin **2-5** (40 µM) was reacted with varying equivalents of EDA (1 to 4) in 10 mM PIPES buffer (pH 7.0) at 25 ℃ for 2 h while stirring and the reaction monitored by UV-vis spectroscopy.

2.7.11 Kinetics of Decomposition of EDA in water

Ethyl diazoacetate (5.31 µM) was stirred at 25 ℃ for 12 h in ddH_2O and the reaction monitored by UV-vis spectroscopy.

2.7.12 Generation of Carbene Complex in D_2O

Ru(CO)TPP-acid (**2-5**) (2.72 mg; 2.96 µmol) was dissolved in 0.7 mL of deuterophosphate buffer in D_2O (20 mM; pD = 7.44 = pH 7.0)[65] and small aliquots of EDA were added resulting in various equivalents and the reaction analyzed by ^{1}H-NMR spectroscopy.

2.7.13 Generation of Carbene Complex in C_6D_6

Ru(CO)TPP (**2-7**) (1.41 mg; 1.90 µmol) was dissolved in 0.7 mL C_6D_6 and small aliquots of EDA were added resulting in 0.5, 1.0 and 3.0 equivalents and the reaction analyzed by ^{1}H-NMR spectroscopy.

2.7.14 Influence of Axial Ligands on Carbene Formation

2-7 (0.73 mg; 0.98 µmol) was dissolved in 0.6 mL C_6D_6 and small aliquots of pyridine, DMAP or mercaptoethanol were added resulting in various equivalents and the reaction analyzed by ^{1}H-NMR spectroscopy. Ethyl diazoacetate (1.0 equivalents) was added in small aliquots.

2.7.15 Examination of the Reaction of Lewis Bases with Ruthenium Carbonyl Porphyrins

2-7 (2.20 mg; 2.96 µmol) was dissolved in 2.1 mL C_6H_6, pyridine was added and the solution stirred at 50 ℃ for 3 h. The solution was analyzed by ESI-MS.

UV-vis spectroscopy: UV-vis absorption spectra were recorded at 25 ℃ u sing a Varian Bio50 UV-vis spectrometer with constant-temperature accessory.

NMR spectroscopy: All NMR spectra (^{1}H, ^{13}C) were recorded on either a Bruker DRX 500 or Mercury VX 400.

2.8 References

(1) Anastas, P. T.; Warner, J. C. *Green chemistry : theory and practice*; Oxford University Press: Oxford [England] ; New York, 1998.

(2) Lindström, U. M. *Organic reactions in water : principles, strategies and applications*; Blackwell Pub.: Oxford ; Ames, Iowa, 2007.

(3) Lu, Y.; Berry, S. M.; Pfister, T. D. *Chemical Reviews* **2001**, *101*, 3047-3080.

(4) Ward, T. R. *Chemistry-a European Journal* **2005**, *11*, 3798-3804.

(5) Wilson, M. E.; Whitesides, G. M. *Journal of the American Chemical Society* **1978**, *100*, 306-307.

(6) Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R. *Journal of the American Chemical Society* **2003**, *125*, 9030-9031.

(7) Klein, G.; Humbert, N.; Gradinaru, J.; Ivanova, A.; Gilardoni, F.; Rusbandi, U. E.; Ward, T. R. *Angewandte Chemie-International Edition* **2005**, *44*, 7764-7767.

(8) Letondor, C.; Humbert, N.; Ward, T. R. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 4683-4687.

(9) Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R. *Journal of the American Chemical Society* **2006**, *128*, 8320-8328.

(10) Letondor, C.; Ward, T. R. *Chembiochem* **2006**, *7*, 1845-1852.

(11) Loosli, A.; Rusbandi, U. E.; Gradinaru, J.; Bernauer, K.; Schlaepfer, C. W.; Meyer, M.; Mazurek, S.; Novic, M.; Ward, T. R. *Inorganic Chemistry* **2006**, *45*, 660-668.

(12) Rusbandi, U. E.; Lo, C.; Skander, M.; Ivanova, A.; Creus, M.; Humbert, N.; Ward, T. R. *Advanced Synthesis & Catalysis* **2007**, *349*, 1923-1930.

(13) Rusbandi, U. E.; Skander, M.; Ivanova, A.; Malan, C.; Ward, T. R. *Comptes Rendus Chimie* **2007**, *10*, 678-683.

(14) Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *Journal of the American Chemical Society* **2004**, *126*, 14411-14418.

(15) Hamachi, I.; Shinkai, S. *European Journal of Organic Chemistry* **1999**, 539-549.

(16) Dolphin, D. *The Porphyrins*; Academic Press: New York, 1978.

(17) Teale, F. W. J. *Biochimica Et Biophysica Acta* **1959**, *35*, 543-543.

(18) Yonetani, T. *Journal of Biological Chemistry* **1967**, *242*, 5008-&.

(19) Battersby, A. R. *Nat Prod Rep* **2000**, *17*, 507-26.

(20) Sheldon, R. A. *Metalloporphyrins in catalytic oxidations*; M. Dekker: New York, 1994.

(21) Murahashi, S. i. *Ruthenium in organic synthesis*; Wiley-VCH: Weinheim, 2004.

(22) Che, C. M.; Huang, J. S. *Coordination Chemistry Reviews* **2002**, *231*, 151-164.

(23) Che, C. M.; Huang, J. S.; Lee, F. W.; Li, Y.; Lai, T. S.; Kwong, H. L.; Teng, P. F.; Lee, W. S.; Lo, W. C.; Peng, S. M.; Zhou, Z. Y. *Journal of the American Chemical Society* **2001**, *123*, 4119-4129.

(24) Ferrand, Y.; Daviaud, R.; Le Maux, P.; Simonneaux, G. *Tetrahedron-Asymmetry* **2006**, *17*, 952-960.

(25) Galardon, E.; Le Maux, P.; Simonneaux, G. *Tetrahedron* **2000**, *56*, 615-621.

(26) Galardon, E.; Le Maux, P.; Toupet, L.; Simonneaux, G. *Organometallics* **1998**, *17*, 565-569.

(27) Harada, T.; Wada, S.; Yuge, H.; Miyamoto, T. K. *Acta Crystallographica Section C-Crystal Structure Communications* **2003**, *59*, M37-M39.

(28) Le Maux, P.; Abrunhosa, I.; Berchel, M.; Simonneaux, G.; Gulea, M.; Masson, S. *Tetrahedron-Asymmetry* **2004**, *15*, 2569-2573.

(29) Le Maux, P.; Juillard, S.; Simonneaux, G. *Synthesis-Stuttgart* **2006**, 1701-1704.

(30) Le Maux, P.; Roisnel, T.; Nicolas, I.; Simonneaux, G. *Organometallics* **2008**, *27*, 3037-3042.

(31) Nicolas, I.; Le Maux, P.; Simonneaux, G. *Tetrahedron Letters* **2008**, *49*, 2111-2113.

(32) Simonneaux, G.; Le Maux, P. *Coordination Chemistry Reviews* **2002**, *228*, 43-60.

(33) Hamaker, C. G.; Mirafzal, G. A.; Woo, L. K. *Organometallics* **2001**, *20*, 5171-5176.

(34) Wolf, J. R.; Hamaker, C. G.; Djukic, J. P.; Kodadek, T.; Woo, L. K. *Journal of the American Chemical Society* **1995**, *117*, 9194-9199.

(35) Dunford, H. B. *Heme peroxidases*; John Wiley: New York, 1999.

(36) Tuynman, A.; Spelberg, J. L.; Kooter, I. M.; Schoemaker, H. E.; Wever, R. *Journal of Biological Chemistry* **2000**, *275*, 3025-3030.

(37) van de Velde, F.; van Rantwijk, F.; Sheldon, R. A. *Trends in Biotechnology* **2001**, *19*, 73-80.

(38) Grieco, P. A. *Organic synthesis in water*; 1st ed.; Blackie Academic & Professional: London ; New York, 1998.

(39) Collman, J. P.; Brothers, P. J.; McElweewhite, L.; Rose, E.; Wright, L. J. *Journal of the American Chemical Society* **1985**, *107*, 4570-4571.

(40) Dattagup.N; Bardos, T. J. *Journal of Heterocyclic Chemistry* **1966**, *3*, 495-&.

(41) Rillema, D. P.; Nagle, J. K.; Barringer, L. F.; Meyer, T. J. *Journal of the American Chemical Society* **1981**, *103*, 56-62.

(42) Nimri, S.; Keinan, E. *Journal of the American Chemical Society* **1999**, *121*, 8978-8982.

(43) Maas, G. *Chemical Society Reviews* **2004**, *33*, 183-190.

(44) Doyle, M. P.; McKervey, M. A.; Ye, T. *Modern catalytic methods for organic synthesis with diazo compounds : from cyclopropanes to ylides*; Wiley: New York, 1998.

(45) Collman, J. P.; Rose, E.; Venburg, G. D. *Journal of the Chemical Society-Chemical Communications* **1993**, 934-935.

(46) Nicolas, I.; Le Maux, P.; Simonneaux, G. *Coordination Chemistry Reviews* **2008**, *252*, 727-735.

(47) Schils, R.; Simal, F.; Demonceau, A.; Noels, A. F.; Eremenko, I. L.; Sidorov, A. A.; Nefedov, S. E. *Tetrahedron Letters* **1998**, *39*, 7849-7852.

(48) Hayashi, T.; Kato, T.; Kaneko, T.; Asai, T.; Ogoshi, H. *Journal of Organometallic Chemistry* **1994**, *473*, 323-327.

(49) Simal, F.; Demonceau, A.; Noels, A. F. *Tetrahedron Letters* **1999**, *40*, 63-66.

(50) Iwasa, S.; Takezawa, F.; Tuchiya, Y.; Nishiyama, H. *Chemical Communications* **2001**, 59-60.

(51) data generated by Tamara C. Bumpers

(52) Galardon, E.; LeMaux, P.; Simonneaux, G. *Chemical Communications* **1997**, 927-928.

(53) Lo, W. C.; Che, C. M.; Cheng, K. F.; Mak, T. C. W. *Chemical Communications* **1997**, 1205-1206.

(54) Regitz, M.; Maas, G. *Diazo compounds : properties and synthesis*; Academic Press: Orlando, 1986.

(55) Staudinger, H.; Gaule, A. *Berichte Der Deutschen Chemischen Gesellschaft* **1916**, *49*, 1897-1918.

(56) Roberts, J. D.; Regan, C. M.; Allen, I. *Journal of the American Chemical Society* **1952**, *74*, 3679-3683.

(57) Bernasconi, C. F.; Leyes, A. E.; Ragains, M. L.; Shi, Y.; Wang, H.; Wulff, W. D. *Journal of the American Chemical Society* **1998**, *120*, 8632-8639.

(58) Gandler, J. R.; Bernasconi, C. F. *Organometallics* **1989**, *8*, 2282-2284.

(59) Herrmann, W. A. *Journal of Organometallic Chemistry* **1975**, *84*, C25-C27.

(60) Schramm, K. D.; Ibers, J. A. *Inorganic Chemistry* **1980**, *19*, 2441-2448.

(61) Penoni, A.; Wanke, R.; Tollari, S.; Gallo, E.; Musella, D.; Ragaini, F.; Demartin, F.; Cenini, S. *European Journal of Inorganic Chemistry* **2003**, 1452-1460.

(62) Latosgrazynski, L.; Cheng, R. J.; Lamar, G. N.; Balch, A. L. *Journal of the American Chemical Society* **1981**, *103*, 4270-4272.

(63) Tatsumi, K.; Hoffmann, R. *Inorganic Chemistry* **1981**, *20*, 3771-3784.

(64) Doyle, M. P.; Dorow, R. L.; Buhro, W. E.; Griffin, J. H.; Tamblyn, W. H.; Trudell, M. L. *Organometallics* **1984**, *3*, 44-52.

(65) Krezel, A.; Bal, W. *Journal of Inorganic Biochemistry* **2004**, *98*, 161-166.

Chapter III

CHARACTERIZATION OF RECONSTITUTED SEMI-SYNTHETIC PROTEINS AS CYCLOPROPANATION CATALYSTS

3.1 Introduction

A critical step in designing a semi-synthetic protein is the choice of a suitable protein scaffold that provides the desired environment for achieving high enantioselectivities. Among possible candidates heme proteins stand out as suitable scaffold because their natural cofactor (heme) can often be readily removed and reconstitution of the resulting apoprotein usually leads to catalytically active holoproteins that are able to catalyze reactions, such as sulfoxidations, stereoselectively (chapter 1.1.6 and chapter 2.1.1).[1,2] That is why a heme protein and its corresponding ruthenium organometallic cofactor had to be chosen that were suitable for reconstitution yielding a semi-synthetic cyclopropanation catalyst.

3.1.1 Selection of Heme Protein

The ideal heme protein for generating a catalytically active semi-synthetic protein has to fulfill several criteria. It should be cheap and readily available in order to carry out cyclopropanation reactions with catalyst loads and amounts of reactants similar to the model reactions with ruthenium porphyrin **2-5** (chapter 2.3.3). Furthermore, the protein

has to be expressable in *E. coli* with good yields in order to allow the optimization of the activity and stereoselectivity of the semi-synthetic protein by mutation of the active site amino acid residues. Moreover, the stability of the protein is a crucial parameter for its viability as a semi-synthetic cyclopropanase. The cofactor removal-reconstitution process requires harsh conditions (low pH, 50 % organic solvent for cofactor extraction). High losses caused by lack of stability under these conditions would render the protein unsuitable. Additionally, it has to be very stable under the chosen reaction conditions for the cyclopropanation reaction. Even partial unfolding could lead to the release of the cofactor, which is already catalytically active. This would result in the decrease or loss of enantioselectivity. The prosthetic group has to be readily exchangeable otherwise the ruthenium porphyrin cofactor cannot be inserted into the active site of the heme protein. Besides, the active-site pocket has to be accessible for the reactants styrene and ethyl diazoacetate. Another imperative requirement is the size of the active-site binding pocket. The binding pocket has to be large enough to accommodate the sterically demanding carbene complex and the olefin in order to facilitate the reaction. Due to the similarities between the epoxidation and cyclopropanation reaction (see chapter 2.1) a good protein expoxidation catalyst can more likely be converted to a good cyclopropanation catalyst.

The epoxidation of olefins in the presence of H_2O_2 is a difficult reaction. A lot of times peroxidase-catalyzed epoxidation reactions exhibited low yields and low enantioselectivity.[3] By far the best heme protein epoxidation catalyst is chloroperoxidase (CPO) from *Caldariomyces fumago*. Styrene derivatives were oxidized to the corresponding (R)-styrene oxides with up to 96% ee.[4,5] However, the reconstitution of CPO is very difficult and proceeds with low efficiency.[6] On the other hand myoglobin is a

well-studied heme protein that was utilized for the generation of catalytically active semi-synthetic proteins (chapter 1.1.6). Myoglobin is closely structurally related to peroxidases and can be converted into an efficient peroxidase with high epoxidation activity. Despite the low epoxidation activity in its native state myoglobin represents a good starting point for the development of a catalytically active cyclopropanase. It meets all important criteria and horse heart myoglobin is even available in gram quantities for a low price. Myoglobin has a fairly open topology, which might be a disadvantage for achieving high enantioselectivities (Figure 3-1).[7] Cytochrome P450 BM3, a heme containing enzyme that was successfully overexpressed in *E. coli*, contains a less open-faced active site pocket and represents therefore a promising candidate for conversion to an active cyclopropanase.[8-11] Moreover, the open side is above the porphyrin ring, thus guiding the approach of the reactant to the reactive metal center in a different fashion. This might prompt a different stereoselectivity and activity in case of a semi-synthetic cytochrome P450.

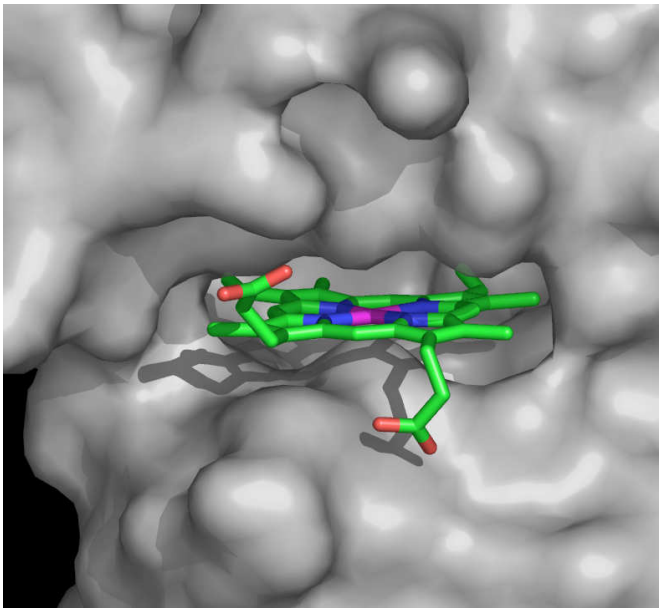


Figure 3-1. Active site topology of myoglobin (protein crystallographic database entry 1AZI)

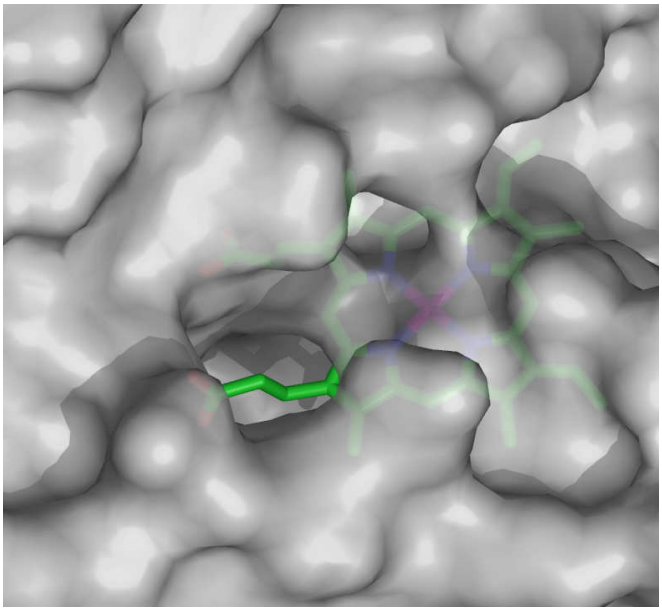


Figure 3-2. Active site topology of cytochrome P450 BM3 (protein crystallographic database entry 1BU7)

3.1.2 Selection of Synthetic Cofactor

Ru(CO)TPP-acid (**2-5**) was used in the studies of cyclopropanation reactions in water due to its simplicity and bulky structure that inhibits dimerization. The active site of myoglobin and cytochrome P450 was evolutionary developed to accommodate its native prosthetic group, which is heme b (iron protoporphyrin IX). Consequently, the ruthenium porphyrin prosthetic group has to be as structurally close to heme b as possible. The obvious choice is ruthenium carbonyl protoporphyrin IX (Ru(CO)PPIX) (Figure 3-3).

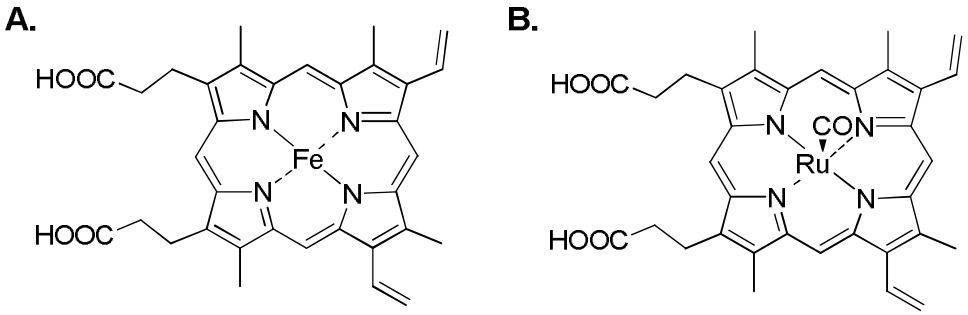
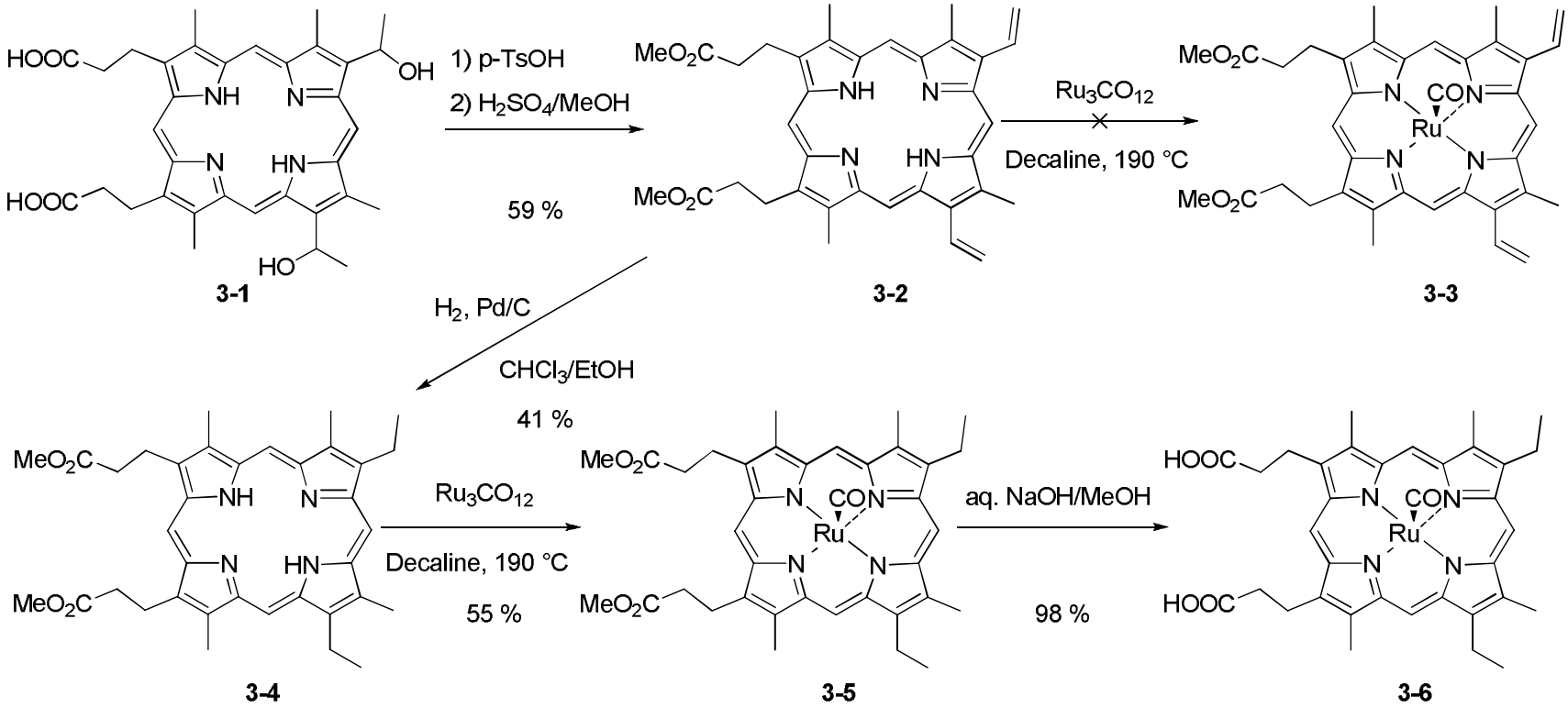


Figure 3-3. A. heme b **B.** Ru(CO)PPIX

3.2 Synthesis of Cofactor

The synthesis of the ruthenium porphyrin cofactor commenced with the dehydration of the commercially available haematoporphyrin (**3-1**) by *p*-toluenesulfonic acid forming two new double bonds (Scheme 3-1). This reaction was followed by esterification of the carboxylic acid moieties yielding protoporphyrin IX dimethylester (**3-2**).[12] The insertion of the ruthenium carbonyl moiety into **3-2** by reaction with $Ru_3(CO)_{12}$ failed probably due to the required harsh conditions of the reaction (high temperatures of about 190 ℃). [13] Ruthenium probably reacted with the two double bonds that are not part of the aromatic system under these conditions, thus leading to the decomposition of

protoporphyrin IX rather than “clean” insertion of the ruthenium metal center. These results are in agreement with Morishima et al. who were also not able to obtain ruthenium carbonyl porphyrin **3-3**.[14] Due to the instability of protoporphyrin IX under the harsh conditions required for the synthesis of many heme analogues, the closely structurally related mesoporphyrin IX complexes were traditionally utilized for studies of reconstituted heme proteins and heme enzymes.[15-20] Therefore **3-2** was partially hydrogenated to give dimethyl mesoporphyrin IX (**3-4**) with palladium on carbon as catalyst.[21] Some over-reduction occurred, which is typical for this type of reaction, yielding a chlorine-type compound (porphyrin with aromatic double bonds, which were partially reduced) that exhibits a UV-vis absorption band at 650 nm. Recently, a straightforward and superior way to hydrogenate **3-2** without the formation of chlorine-type derivatives was reported (after the studies had been finished).[22] Ruthenium porphyrin **3-5** was obtained by insertion of the ruthenium carbonyl moiety into **3-4** with $Ru_3(CO)_{12}$ under high temperatures (190 ℃). [13] Very recently a new synthetic method to produce **3-5** with higher yields was reported (after the studies had been finished), which utilized a ruthenium-DMF complex instead of $Ru_3(CO)_{12}$.[23] The ruthenium porphyrin ester (**3-5**) was hydrolyzed to the final water-soluble ruthenium(II)carbonyl mesoporphyrin IX **3-6** (Ru(CO)MPIX).


HOOC
OH
HOOC
HO
1) p-TsOH
2) H2SO4/MeOH
59 %
3-1
MeO2C
MeO2C
3-2
Ru3CO12
Decaline, 190 °C
MeO2C
MeO2C
Ru
3-3
H2, Pd/C
CHCl3/EtOH
41 %
MeO2C
MeO2C
Ru3CO12
Decaline, 190 °C
55 %
3-4
MeO2C
MeO2C
Ru
3-5
aq. NaOH/MeOH
98 %
HOOC
HOOC
Ru
3-6

Scheme 3-1.

3.3 Characterization of Native Myoglobin Reconstituted with Ru(CO)MPIX

3.3.1 Reconstitution of Native Myoglobin

In order to incorporate the synthetic cofactor the previously described cofactor reconstitution method based on non-covalent attachment (chapter 1.1.6) was utilized (Figure 3-4).[1] The heme group of myoglobin can be readily removed by the acido-butanone method.[24,25] Acidification (pH ≈ 2) leads to protonation of amino acid residues in the active site and thus interrupts stabilizing salt bridges to the carboxylic acids of the heme. This results in the loss of the heme group, which can then be extracted with butanone. The decrease of helix content due to the loss of heme is only 20%, as determined by circular dichroism (CD), which means that backbone conformation and basic folding features are well preserved.[26,27] Reconstitution and full recovery of the helix content and activity is achieved simply by changing the pH to around neutral and addition of the cofactor. The reconstitution of apomyoglobin with the natural heme b is surprisingly fast. It has been shown by ESI-MS that the reconstitution of myoglobin can be complete in as little as 9 s.[28] Since reconstitution is mostly governed by hydrophobic interactions and hydrogen-bonding and not interactions with the metal center, the reconstitution with the structurally very similar ruthenium mesoporphyrin was expected to proceed in a similar fashion. Indeed, addition of apomyglobin to **3-6** resulted quickly in spectral shifts that are characteristic for reconstituted myoglobin (Figure 3-5).[18,19] The absorption maximum (soret band) is red-shifted by 10 nm compared to “free” **3-6**. The splitting of one of the bands (548 nm) in the visible range of **3-6** into two bands (554 nm and 546 nm) of the reconstituted myoglobin (Ru(CO)Mb) supposedly derives from

lowering of the symmetry upon incorporation into the active site due to steric hindrance (Figure 3-6).

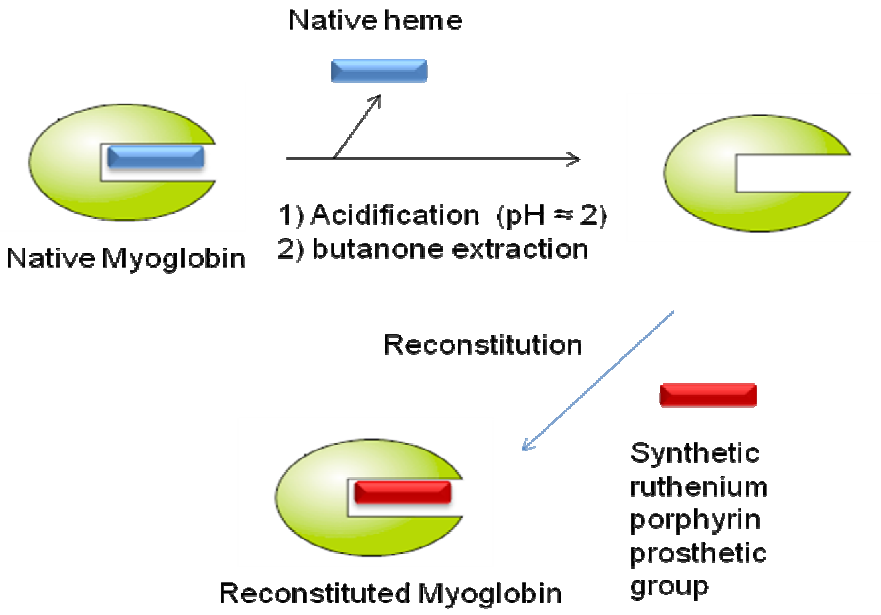


Figure 3-4. Reconstitution of Myoglobin with synthetic cofactor

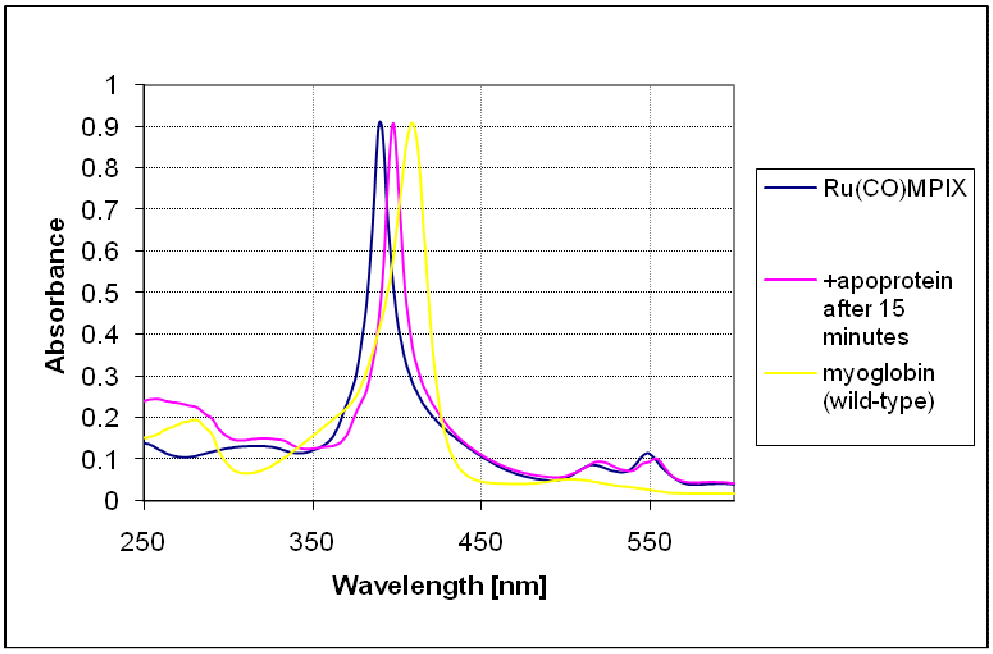


Figure 3-5. UV-vis spectra of myoglobin reconstituted with **3-6** (lilac) in comparison with wild-type myoglobin (yellow) and free cofactor **3-6** (blue) (spectra normalized to soret band).

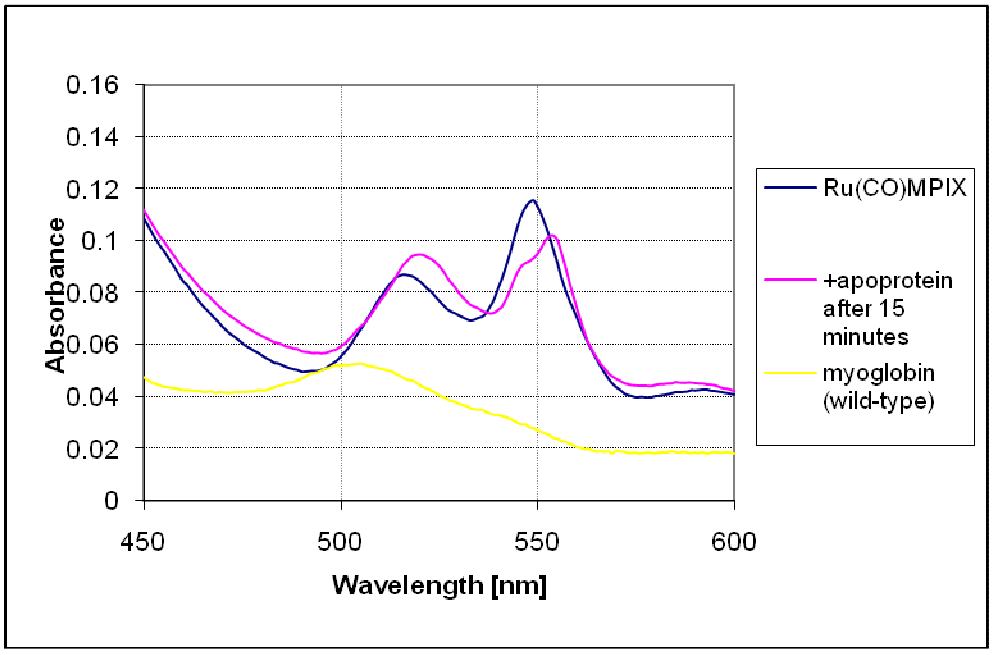
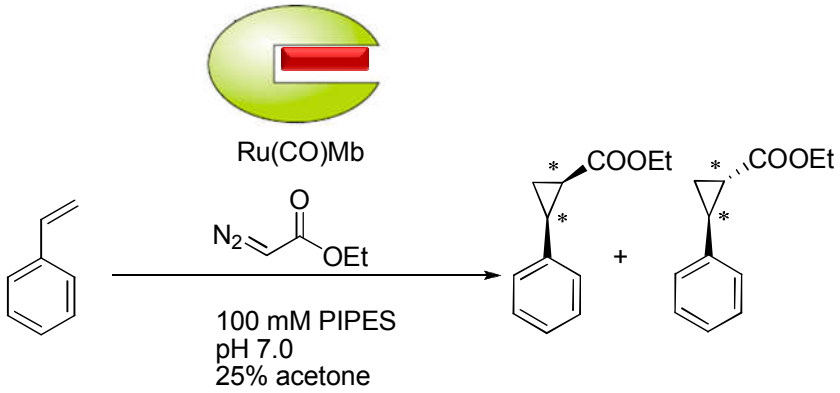


Figure 3-6. UV-vis spectra of myoglobin reconstituted with **3-6** (lilac) in comparison with wild-type myoglobin (yellow) and free cofactor **3-6** (blue) with focus on the visible range (spectra normalized to soret band).

3.3.2 Cyclopropanation Using Myoglobin Reconstituted with Ru(CO)MPIX as Catalyst

The cyclopropanation reaction of styrene with ethyldiazoacetate catalyzed by the semi-synthetic protein Ru(CO)Mb (apomyoglobin reconstituted with **3-6**) was carried out in aqueous medium (Scheme 3-2). Unfortunately, the yield was very low (about 1 %) and racemic. Due to the very high optical activity of the expected cyclopropanation products even small enantiomeric excess would have been detected by polarimetry.[29,30] The cyclopropanation products were proably formed by small amounts of released **3-6** because of decomposition or partial denaturation of Ru(CO)Mb.



Scheme 3-2.

The difference between the cyclopropanation reaction with **3-6** and Ru(CO)Mb is that in case of Ru(CO)Mb, **3-6** is embedded in the active site of the protein. Since 3-6 shows catalytic activity in the absence of protein, the inhibition of the catalyst should be due to active site. Steric crowdedness or binding of axial ligands to the embedded cofactor may be the reason for the inhibition of the cyclopropanation reaction (Figure 3-7). The

proximal histidine 93 (beneath the porphyrin) binds to the metal center of the prosthetic group. It has been previously shown (NMR studies chapter 2.5.3) that Lewis base axial ligands can deactivate the ruthenium porphyrin catalyst. Histidine contains an imidazole moiety that might inhibit the reaction with EDA.

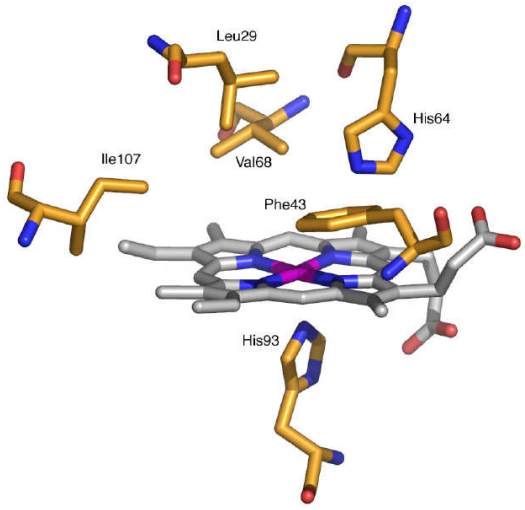


Figure 3-7. Cartoon representation of active site of horse heart myoglobin (entry 1AZI from protein crystallographic database)

3.3.3 The Influence of Imidazole on the Catalyst Activity of Ru(CO)MPIX

3.3.3.1 Analysis by UV-vis Spectroscopy

The reaction of Ru(CO)MPIX (**3-6**) with EDA in aqueous medium in the presence and absence of imidazole were conducted in order to model the influence of the axially bound histidine in Ru(CO)Mb. First, the reactivity of **3-6** with regard to EDA was tested and monitored by UV-vis spectroscopy (Figure 3-8). The UV-vis spectrum, which is almost identical to the spectrum of the reaction of Ru(CO)TPP-acid (**2-5**) with EDA

(chapter 2.4.2; Figure 2-12), confirmed rapid reaction with EDA. However, basically no spectral changes could be observed upon addition of EDA if imidazole is present in the reaction mixture. These results suggest that the imidazole moiety of histidine 93 that is bound to the metal center of the cofactor (Figure 3-7) could indeed inhibit the cyclopropanation activity of the semi-synthetic protein Ru(CO)Mb.

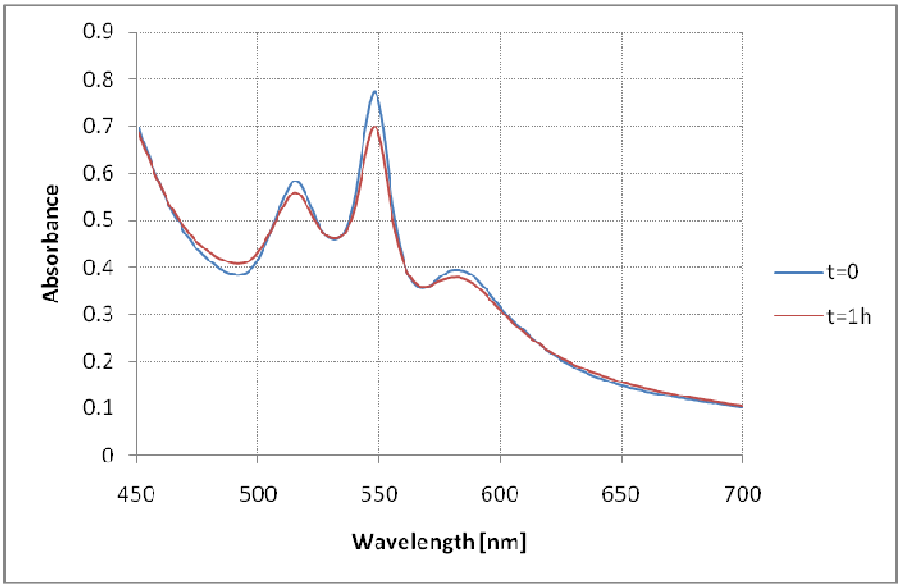


Figure 3-8. UV-vis spectrum of the reaction of **3-6** with EDA in PIPES buffer at pH 7.0

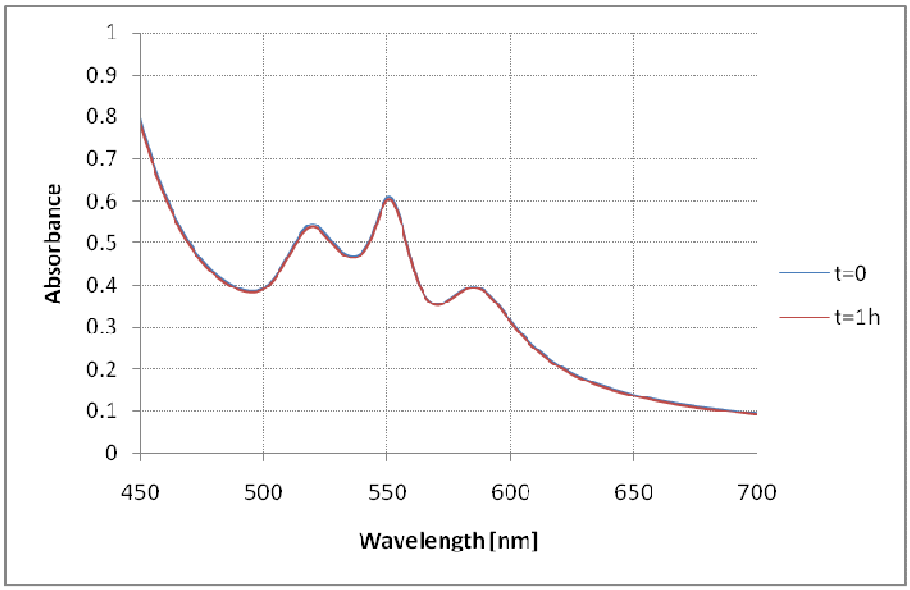
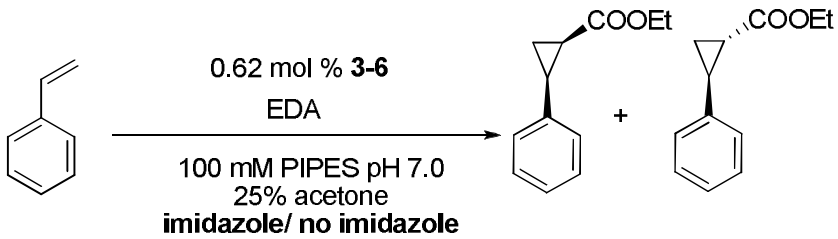


Figure 3-9. UV-vis spectrum of the reaction of **3-6** with EDA in presence of imidazole (30 mM) in PIPES buffer at pH 7.0

3.3.3.2 Analysis by Gas Chromatography

The cyclopropanation of styrene with EDA by Ru(CO)MPIX (**3-6**) in aqueous medium in the presence and absence of imidazole were carried out to confirm the deactivating effect of imidazole suggested by reactions monitored by UV-vis spectroscopy (Figure 3-8 & 3-9). Analysis by GC-FID illustrated the inhibitory effects of an axially bound ligand such as imidazole (Figure 3-10). Only minuscule amounts of cyclopropanation products were found in case of imidazole in the reaction mixture, in striking contrast to the reaction without imidazole.



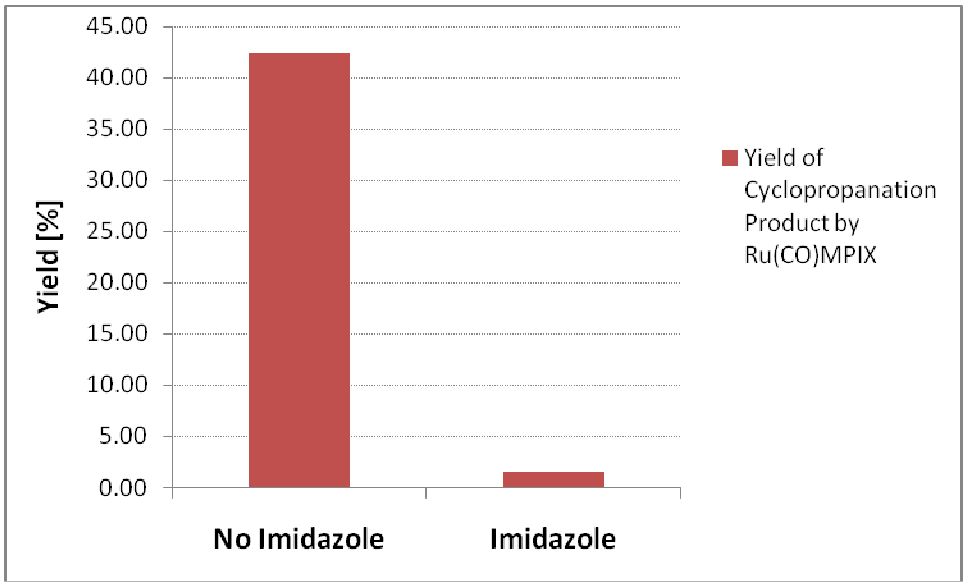
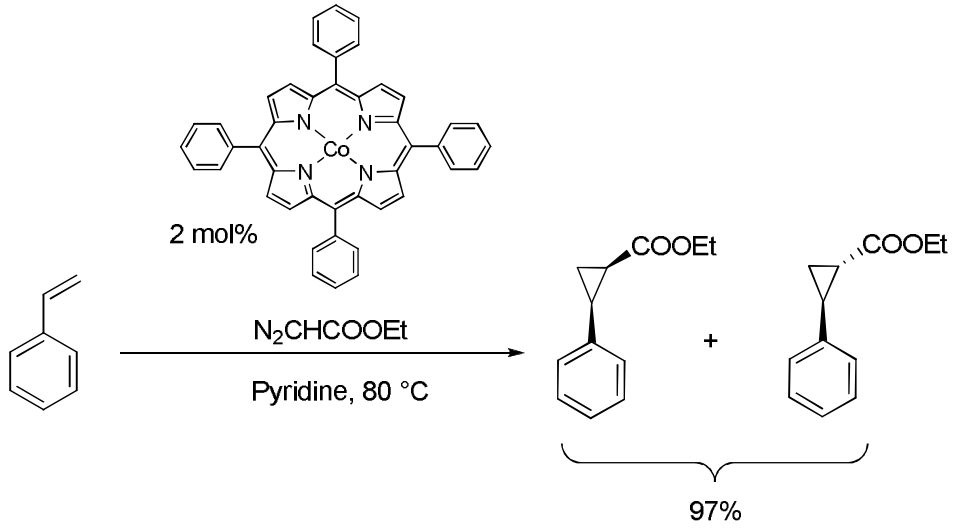


Figure 3-10. Analysis (GC-FID) of cyclopropanation of styrene by **3-6** with or without imidazole

3.4 Cobalt Porphyrins as Alternatives to Ruthenium Porphyrins

3.4.1 Selection of Cobalt Porphyrins as Cyclopropanation Catalysts

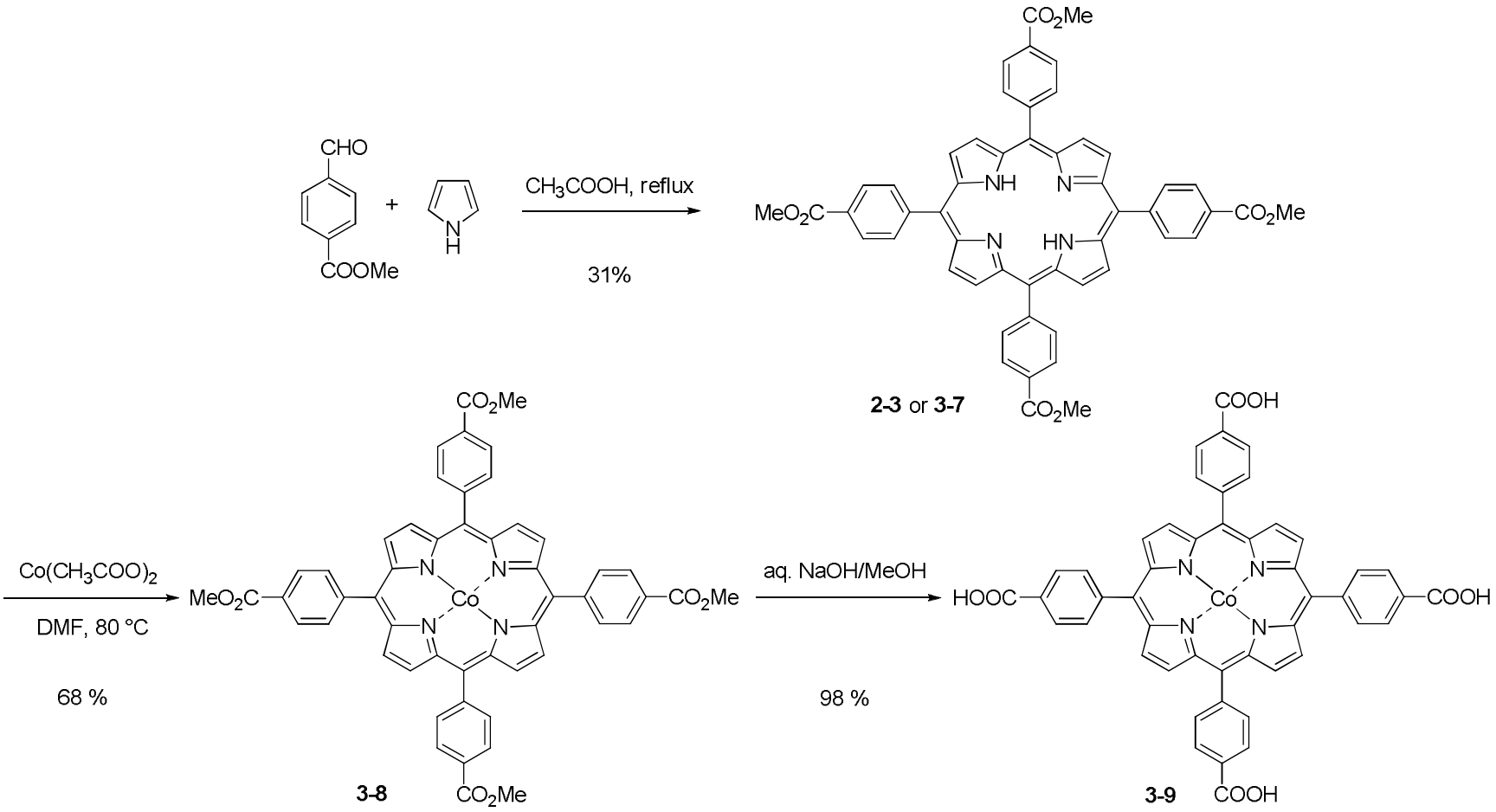
Cobalt porphyrins are able to catalyze the cyclopropanation of styrene with EDA in pyridine as a solvent (Scheme 3-3).[31] Due to the similarities of pyridine and imidazole in the deactivation behavior cobalt porphyrins might not get deactivated by imidazole, thus representing a viable alternative type of cofactor for the active site of myoglobin.



Scheme 3-3.

3.4.2 Synthesis of a Water-Soluble Cobalt Porphyrin

A water-soluble cobalt porphyrin was synthesized utilizing the same strategies as for the synthesis of the water-soluble ruthenium porphyrin **2-5** (chapter 2.2) (Scheme 3-4). The cobalt metal center was inserted into **3-7** via reaction with cobalt(II) acetate in DMF.[32] The synthesis was concluded by the hydrolysis of the four ester moieties yielding the water-soluble cobalt(II) *meso*-tetrakis-(4-carboxyphenyl)porphyrin (**3-9**).


CHO
COOMe
+
N
H
CH3COOH, reflux
31%
CO2Me
MeO2C
CO2Me
NH
N
N
HN
CO2Me
2-3 or 3-7
Co(CH3COO)2
DMF, 80 °C
68 %
MeO2C
CO2Me
CO2Me
Co
3-8
aq. NaOH/MeOH
98 %
COOH
HOOC
COOH
Co
3-9
COOH

Scheme 3-4.

3.4.3 Cyclopropanation of Styrene by Cobalt Porphyrins in the Presence of Imidazole

The cyclopropanation of styrene with EDA by cobalt porphyrin **3-9** was carried out in aqueous medium with and without imidazole in order to examine the impact of this axial ligand on the activity of the catalyst. Unfortunately, the overall results of the cobalt porphyrin catalyzed reactions were disappointing (Figure 3-11). Examination of the reaction by GC-FID revealed that the yield of the cyclopropanation with **3-9** was only a fraction of the yield of the same reaction with ruthenium porphyrin **3-6** (in the absence of imidazole). Most importantly, imidazole also deactivated cobalt porphyrin **3-9**. The excess of imidazole in the reaction mixture probably led to the formation of a 1:2 complex (**3-9** : imidazole) because both axial positions can be occupied by a ligand resulting in a six coordinated species. A 1:1 complex might still be catalytically active. However, it is very difficult to carry out the experiments with a 1:1 complex, even if it could be generated reliably. The cobalt porphyrin imidazole complex can dissociate yielding an already catalytically active species. Since no dissociation constants are known it would be very challenging to ensure catalysis with exclusively a 1:1 complex. Besides, cobalt porphyrins are harder to handle than ruthenium carbonyl porphyrins. Cobalt porphyrins bind oxygen tightly, thus inhibiting cyclopropanation reactions. That is why the reactions had to be carried out strictly under argon atmosphere. All in all, cobalt porphyrins were inferior catalysts for the cyclopropanation of styrene in aqueous medium, which also get deactivated by imidazole. Hence, they do not constitute a viable alternative to ruthenium porphyrins.

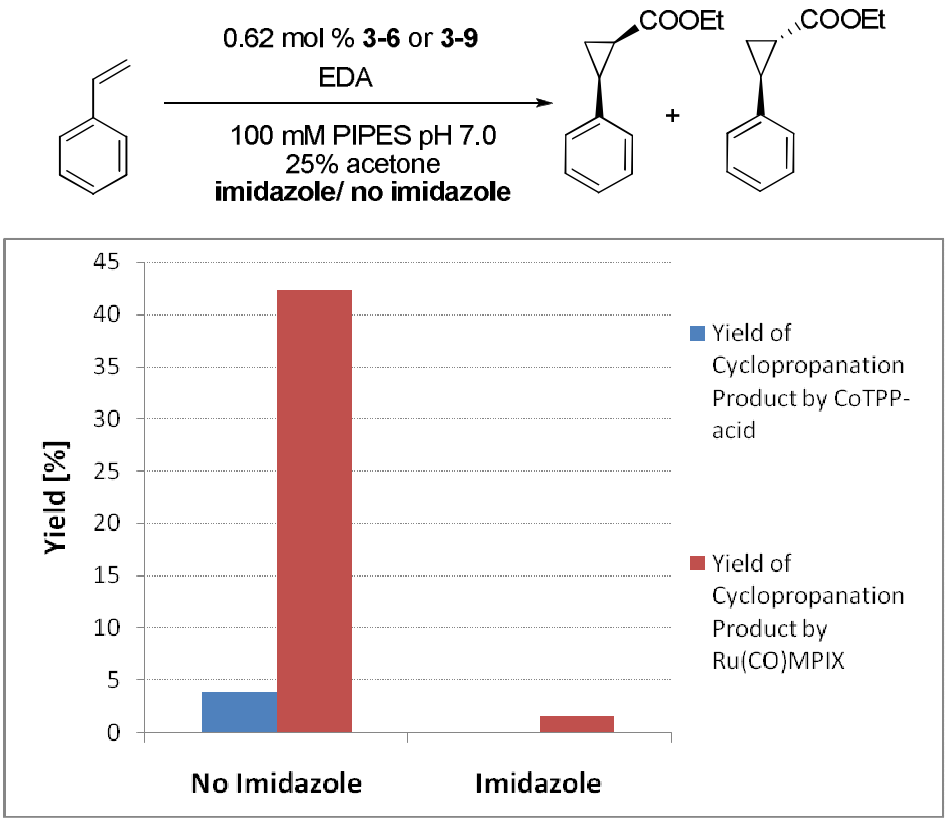


Figure 3-11. Analysis (GC-FID) of cyclopropanation of styrene by **3-6** compared to **3-9** with or without imidazole.

3.5 Characterization of Myoglobin Mutants Reconstituted with Ru(CO)MPIX

3.5.1 Generation of Myoglobin Mutants

Due to failure of ruthenium and cobalt porphyrins to catalyze the cyclopropanation of styrene with imidazole as axial ligands the histidine moieties of the active site of myoglobin had to be mutated (Figure 3-7). Histidine 93 binds directly to the porphyrin metal center and histidine 64 represents a bulky group within the binding pocket that might obstruct the reaction. Glycine is the smallest amino acid and cannot

bind to the ruthenium metal center. Therefore the myoglobin mutants H64G, H93G and H64G/H93G were generated for over-expression in *E.coli* (molecular cloning performed by Dr. Javier Chaparro-Riggers). The myoglobin H93G mutant and the H64D/H93G double mutant have been previously described in the literature.[33,34] The new double mutant H64G/H93G is probably stable and expressable in *E. coli* due to the similarity to myoglobin H64D/H93G.

3.5.2 Properties of Myoglobin Mutants

His-tagged horse-heart myoglobin and myoglobin mutants were overexpressed in *E.coli* and purified (standard His-tag purification) (Figure 3-12). Porphyrins usually show a significant shift of the soret band upon binding of axial ligands. The shift shown by binding of imidazole is very similar to the shift after reconstitution of myoglobin. Interestingly, the mutants did not exhibit any shift at the soret band despite different axial ligands. This is in agreement with published UV-vis spectra of myoglobin mutants in the literature. The purity of the expressed proteins was confirmed by SDS-PAGE and the ratio of the protein concentrations (determined by Bradford assay) to absorption maxima at the soret band were determined. These ratios indicate the relative level of heme incorporation, assuming 100 % incorporation in case of native myoglobin (Figure 3-13). The myoglobin mutant H93G (proximal histidine replaced by glycine) had nearly the same level of heme incorporation as the wild-type, thus showing a remarkable stability despite the absence of anchoring of the heme-group by direct binding of the histidine residue to the porphyrin metal center. In contrast, H64G (distal histidine replaced by glycine) exhibited a lower level of heme incorporation (it also has by far the lowest

expression yield of all myoglobin mutants). The myoglobin double mutant H64G/H93G showed the lowest level of incorporation. However, it was still stable enough to retain more than half of the embedded heme group. This trend of stability is consistent with kinetic studies of the heme loss of various myoglobin mutants.[35]

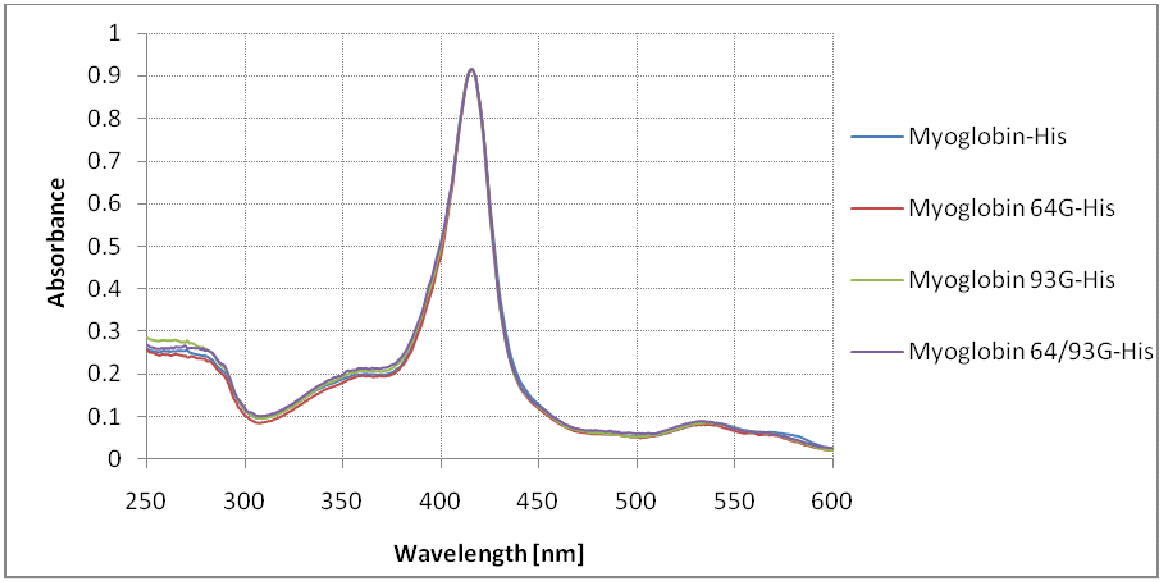


Figure 3-12. Normalized UV-vis spectra of expressed horse-heart myoglobin mutants

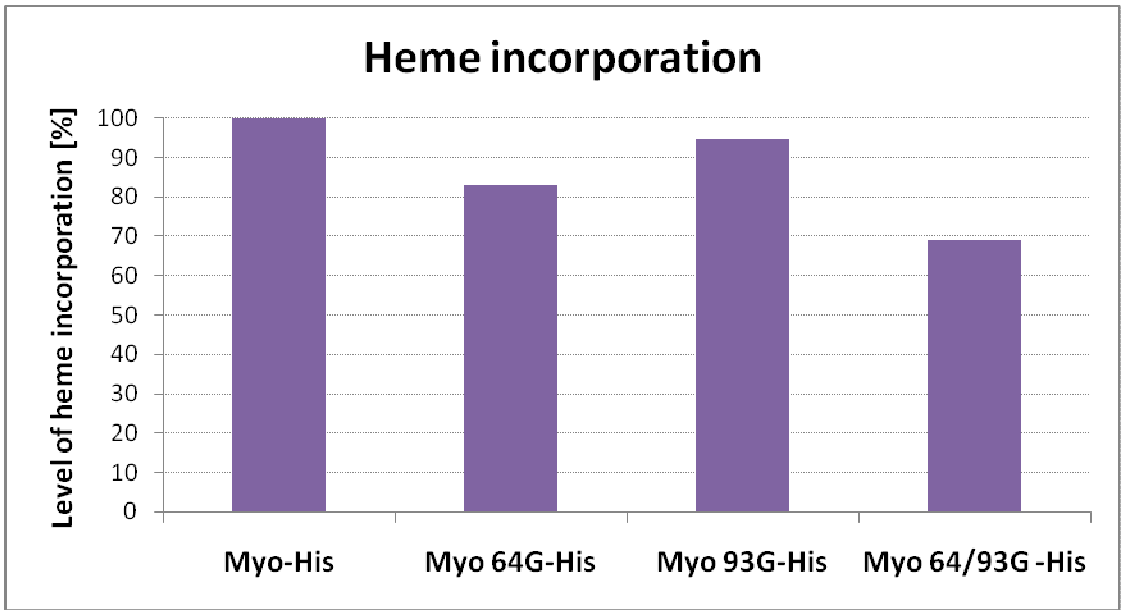


Figure 3-13. The relative level of heme incorporation of myoglobin mutants.

3.5.3 Cyclopropanation with Myoglobin H64G/H93G Reconstituted with Ru(CO)MPIX

The cyclopropanation reaction of styrene with ethyldiazoacetate was catalyzed by the semi-synthetic protein Ru(CO)Mb H64G/H93G (apoprotein reconstituted with **3-6**). Unfortunately, the results of the reaction catalyzed by the ruthenium porphyrin reconstituted double mutant are as those for the myoglobin wild type catalyzed reaction (low yield, racemic mixture of products). These results were confirmed by the reaction of Ru(CO)Mb H64G/H93G with EDA, monitored by UV-vis spectroscopy (Figure 3-14). Almost no change could be observed upon addition of EDA; neither the strong absorption band of EDA at 250 nm nor the soret band showed any signs of reaction. The reason for the slight initial change of the spectrum in the visible range is probably only

due to solvent effects (EDA stock solution was not dissolved in the reaction buffer but in DMF) because the spectrum did not continue to change (Figure 3-15).

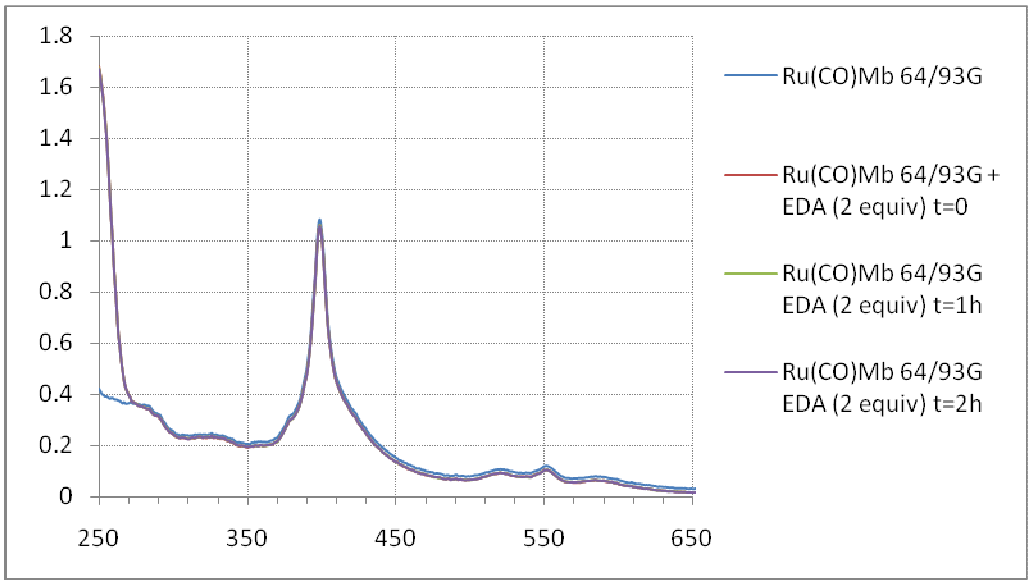


Figure 3-14. UV-vis spectra of reaction of Ru(CO)Mb 64/93G with EDA

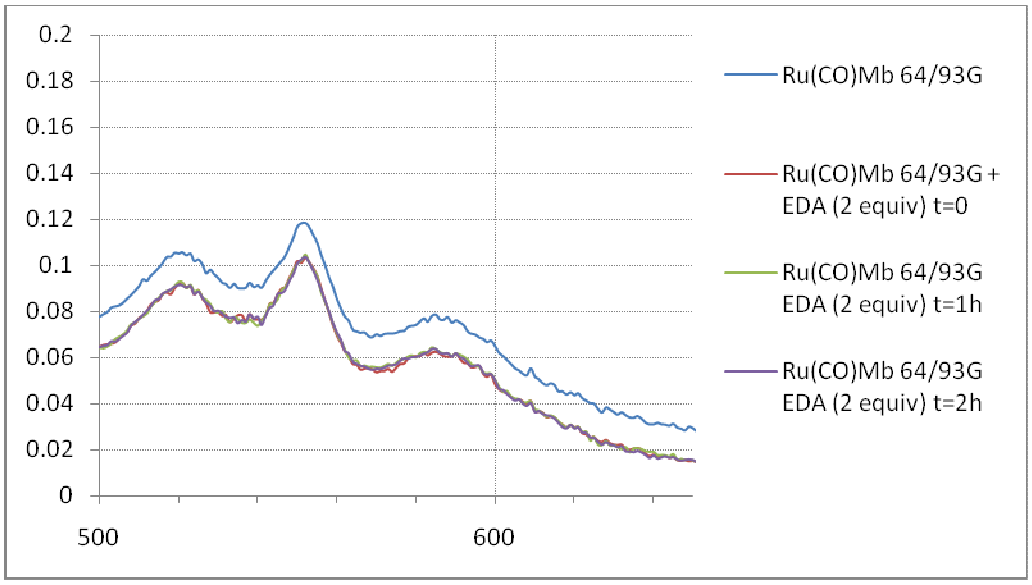


Figure 3-15. UV-vis spectra of reaction of Ru(CO)Mb 64/93G with EDA

3.5.4 Cyclopropanation with Myoglobin H64G/H93G Reconstituted with RuMPIX

Cyclopropanation attempts with myoglobin or myoglobin mutants reconstituted with Ru(CO)MPIX had failed. The initial reaction of ethyl diazoacetate with ruthenium carbonyl porphyrins might occur on the axial position of the metal center opposite of the carbonyl moiety. Since this position is most likely obstructed by the protein backbone the carbonyl moiety was removed before reconstitution, simply by reacting **3-6** with EDA and styrene. Utilizing this method did again not lead to any cyclopropanation products with enantiomeric excess. The failure of this method is illustrated by UV-vis spectra of this reaction (Figure 3-16 + Figure 3-17). Ru(CO)MPIX (**A**) was reacted with EDA and styrene in a mixture of PIPES buffer and acetone for one hour (**B**) in order to remove the carbonyl functionality of **3-6**. Significant spectral changes could be seen despite the lack of clearly defined isosbestic points. Upon addition of apomyoglobin 64/93G immediate characteristic spectral shifts occurred confirming the generation of the semi-synthetic protein (**C**). However, all spectral changes ceased suggesting the inhibition of the catalytic activity (**D**). This clearly shows the inability of ruthenium porpohyrin reconstituted myoglobin mutants to catalyze cyclopropanation reactions.

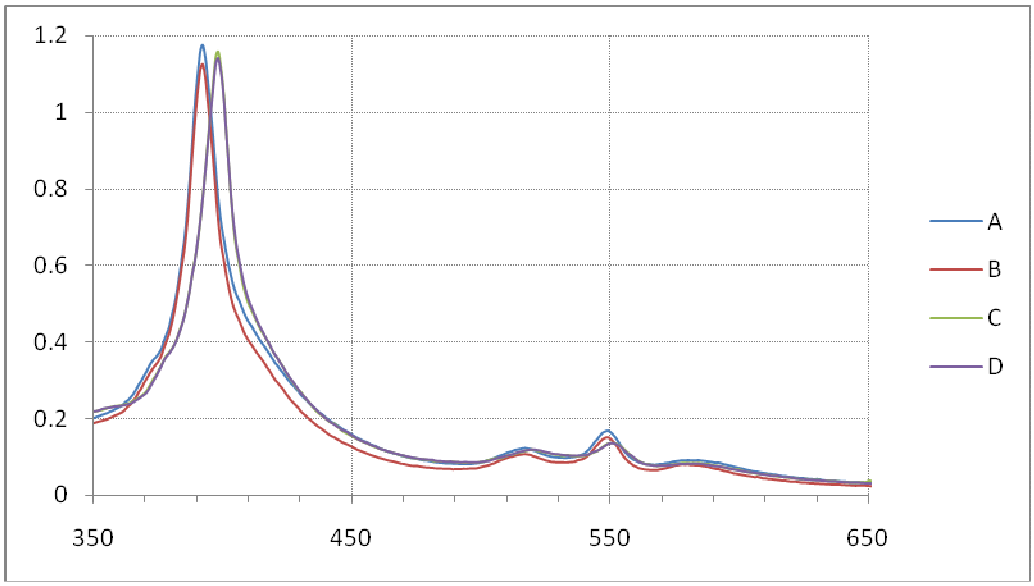


Figure 3-16. Cyclopropanation of styrene by **3-6** and RuMb. **(A). 3-6 (B). 3-6** + EDA + styrene t = 1 h **(C). B** + apomyoglobin 64/93G t = 0 **(D). C** t = 1 h.

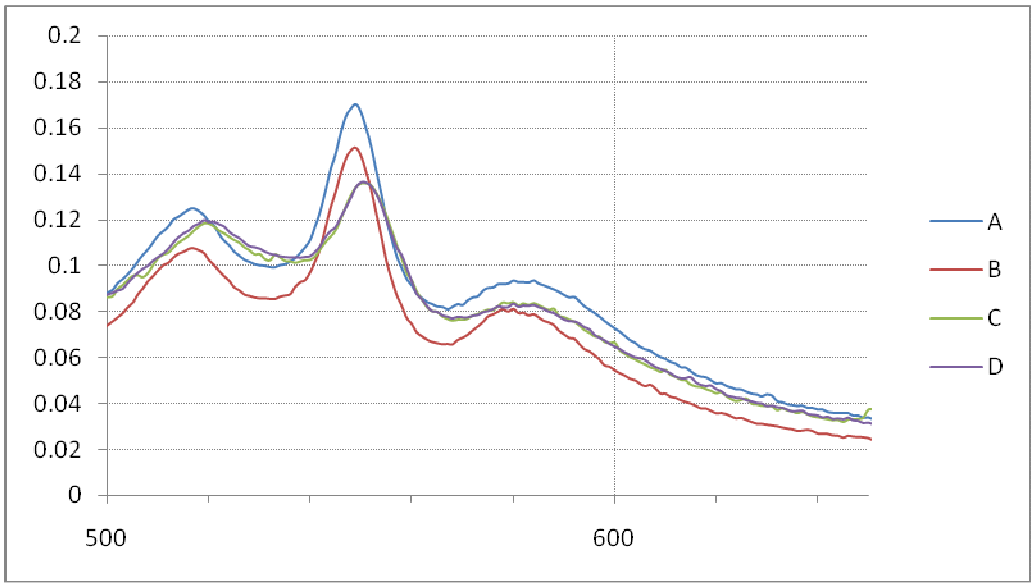


Figure 3-17. Cyclopropanation of styrene by **3-6** and RuMb. **(A). 3-6 (B). 3-6** + EDA + styrene t = 1 h **(C). B** + apomyoglobin 64/93G t = 0 **(D). C** t = 1 h.

3.6 Non-Specific Binding of Synthetic Prosthetic Group

Generally, either equimolar amounts or a slight excess of apomyoglobin was used for reconstitution with **3-6** followed by purification with size-exclusion chromatograpy (gel-filtration). No relative change of the UV-vis spectrum of the semi-synthetic protein solution before and after gel filtration was observed (data not shown). In the literature no significant excess of the cofactor was used for reconstitution. If a significant excess of the cofactor is used then the UV-vis spectrum before and after gel filtration still remains the same (after normalization of the spectra based on protein concentration). Hence, the size-exclusion chromatography did not successfully separate excess unincorporated ruthenium porphyrin. In the absence of protein Ru(CO)MPIX stays on top of the column, which is the expected behavior of a small molecule in size-exclusion chromatography. In the presence of protein even a 20 fold excess of Ru(CO)MPIX was eluted together with the protein. This suggests that non-specific binding is the cause of the failure to separate excess **3-6**. Non-specific interactions are usually reduced by an increase of the ionic strength. However, even a buffered solution with one molar NaCl did not inhibit the “sticking” of Ru(CO)MPIX to the protein. The change of other conditions, such as pH, column loading, column length, etc. did not result in an improvement of the separation.

Ruthenium porphyrin **3-6** contains a carbonyl functionality that exhibits a characteristic C=O stretch in an IR spectrum. This feature can be utilized to study non-specific interactions between **3-6** and myoglobin because binding to the surface would probably cause a shift of the carbonyl stretch. Water very strongly absorbs light in the frequency range of interest for studying proteins. This problem can be circumvented by using Attenuated Total Reflection Infrared Spectroscopy (ATR-IR).[36] Ru(CO)MPIX was

added to solutions of green fluorescent protein (GFP), horse heart myoglobin wild type and expressed (*E.coli*) His-tagged wild-type myoglobin (Figure 3-18). Both types of myoglobin were in the holo-form; therefore no incorporation of **3-6** into the active site should have been possible. GFP has no heme binding sites and was selected as a "control protein". Mixing GFP with Ru(CO)MPIX resulted in a shift compared to "free" Ru(CO)MPIX (see black line in spectrum) suggesting non-specific binding to be origin of this spectral change. Unfortunately, the myoglobin ruthenium carbonyl mixtures could not be resolved enough to clearly compare the IR stretches with each other.

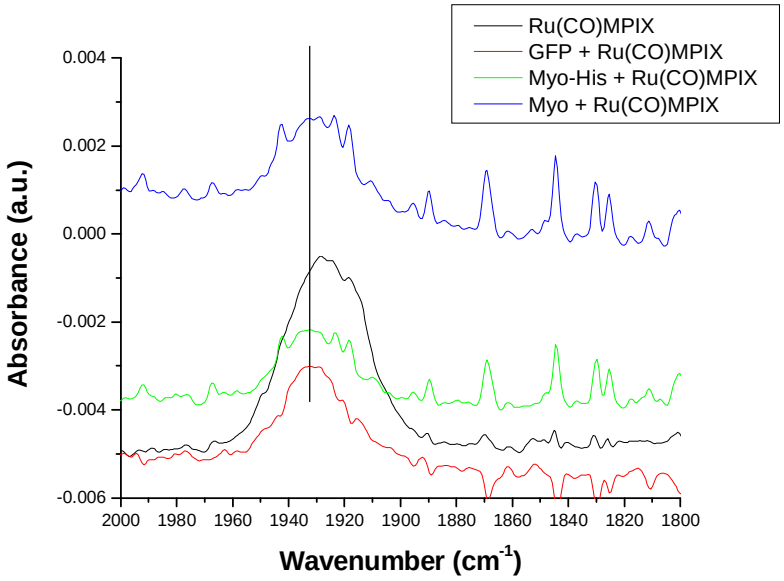


Figure 3-18. ATR-IR spectra of mixtures of **3-6** with myoglobin and GFP.

3.7 Cyclopropanation by Cytochrome P450 154C1

The CYP154C1 C355A mutant (proximal anchoring cysteine replaced by alanine) was overexpressed in *E. coli*. Unfortunately, no heme group was found in purified protein (lack of the typical strong red color; confirmed by UV-vis spectroscopy) suggesting that

the cysteine moiety is crucial for the stability of the heme group in the active site. All attempts to “reconstitute” this protein with Ru(CO)MPIX failed. The cofactor does not get incorporated into the active site and slowly precipitates overtime. Therefore cyclopropanation reactions could not be carried out.

3.8 Cyclopropanation by Anchored Ruthenium Porphyrin - BSA Conjugates

Komatsu et al. reported on the formation of iron tetraphenylporphyrin serum albumin conjugates.[37] Therefore Ru(CO)TPP-acid (**2-5**) was reacted with bovine serum albumin in order to provide a chiral framework for cyclopropanation reactions (Figure 3-19). Ruthenium porphyrin **2-5** exhibited a shift similar to **3-6** after incorporation into myoglobin (**B**). BSA has two porphyrin binding sites, therefore addition of another equivalent of **2-5** did not result in a change of the soret band (**C-D**). Unfortunately, cyclopropanation reactions with ruthenium porphyrin – BSA conjugates proceeded in a similar way like myoglobin reconstituted with **3-6**; small amounts of formed cyclopropanation product and racemic products.

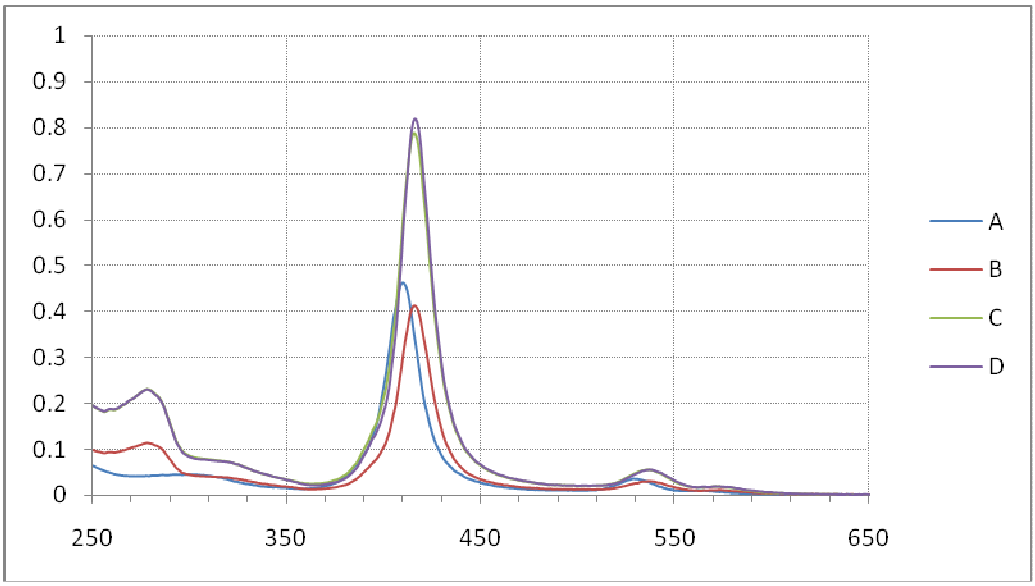


Figure 3-19. Generation of ruthenium porphyrin – BSA attachment. **(A) 2-5**. **(B)** A + BSA (1 equiv). **(C)** B + **2-5** (1 equiv) t = 0. **(D)** C t = 2 h.

3.9 Conclusion

Myoglobin and myoglobin mutants were successfully reconstituted with the heme-like ruthenium carbonyl porphyrin **3-6**. None of the generated semi-synthetic proteins were catalytically active "cyclopropanases". Deactivation of the catalytic activity by axial histidine ligands represented a big obstacle for this project. The origin of the lack of activity in case of missing histidine ligands in the active site is not known. Non-specific binding of the ruthenium cofactor to the protein and the resulting purification problems were obstacles that could never be overcome. Many reactions have been catalyzed by semi-synthetic proteins but never cyclopropanation reactions. This is an indication for the difficulty of the reaction with semi-synthetic proteins. Apparently, converting myoglobin into an active cyclopropanase requires more drastic changes then a few

simple point mutations of the active site. This is a challenging task, which was not further pursued.

3.10 Experimental Section

3.10.1 Materials and Reagents and General Techniques

All starting materials, reagents and solvents were purchased from Aldrich (Milwaukee, WI), Fisher Scientific (Pittsburgh, PA) or VWR (West Chester, PA). NMR: δ in ppm versus SiMe4 (0 ppm, ^{1}H, 400 MHz). MS: selected peaks; m/z. Flash chromatography: Merck silica gel (240-400 mesh). TLC: 0.25 mm, Merck silica gel 60 F_{254}, visualizing at 254 nm or with permanganate solution (2 % $KMnO_4$, 5 % K_2CO_3). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). Ampicillin was used at 50 µg/mL and Kanamycin at 30 µg/mL concentration. Unless otherwise stated, standard protocols were used. All plasmids were verified by DNA sequencing. Mutated sites are in indicated in bold and red.

3.10.2 Synthesis

Protoporphyrin IX dimethylester (**3-2**) was synthesized following the published procedure.[12]

Attempted synthesis of ruthenium(II)carbonyl protoporphyrin IX dimethylester (3-3). Modifications of the procedure of Rillema et al. were used.[13] A mixture of **3-2** (50 mg, 84.6 µmol) and Ru_3CO_{12} (54 mg, 84.5 µmol) in 10 mL of decaline were heated for 24 hours under reflux and then allowed to cool to room temperature. No starting material or desired product could be found in the reaction mixture.

Mesoporphyrin IX dimethylester (3-4). Modifications of the procedure of Muir et al. were used.[21] A solution of protoporphyrin IX dimethylester **3-2** (750 mg, 1.27 mmol) in 1,2-dichloroethane (32 mL) and ethanol (50 mL) was hydrogenated at room temperature and ambient pressure in the presence of Pd on activated carbon as catalyst (78 mg, 5 wt % Pd, 3 mol %) for three days. The crude product was purified by using flash chromatography on silica gel ($CHCl_3$/EtOAc 50:1) providing **3-4** as a purple powder (310 mg, 41 %). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 10.10, 10.09 (2s, 4H); 4.42 (t, J = 7.6 Hz, 4H); 4.09 (q, J = 7.6, 4H); 3.67, 3.65 (2s, 6H); 3.64, 3.63 (2s, 12H); 3.28 (t, J = 7.6, 4H); 1.85 (t, J = 7.62, 6H); -3.76 (brs, 2H). Spectroscopic data are in agreement with those reported.[22]

Ruthenium(II)carbonyl mesoporphyrin IX dimethylester (3-5). Modifications of the procedure of Rillema et al. were used.[13] A mixture of **3-4** (285 mg, 0.48 mmol) and Ru_3CO_{12} (291.4 mg, 0.46 mmol) in 50 mL of decaline were heated for 5 hours under reflux and then allowed to cool to room temperature. The suspension was filtered and the remaining solid extracted several times with a 100 mL CH_2Cl_2-acetone mixture (95:5). The combined extracts were concentrated under reduced pressure and the crude product was purified by using flash chromatography on silica gel ($CHCl_3$/EtOAc 50:1, gradually changed to $CHCl_3$/EtOAc 15:1) providing **3-5** as a purple powder (190 mg, 55 %). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 9.96, 9.95, 9.94 (3s, 3H); 9.85 (s, 1H); 4.33-4.26 (m, 4H); 4.02 (q, J = 7.6 Hz, 4H); 3.65, 3.62 (2s, 6H); 3.59, 3.58, 3.57, 3.56 (4s, 12H); 3.27-3.20 (m, 4H); 1.88 (t, J = 7.6 Hz, 6H). FAB-MS: *m/z* 723.2 [M+H$^+$]. FAB-HRMS calculated for [M+H$^+$] C_{37} H_{41} N_4 O_5 723.2120, found 774.2063.

Ruthenium(II)carbonyl mesoporphyrin IX (3-6). Modifications of the procedure of Nimri et al. were used.[38] Ruthenium porphyrin **3-5** (43 mg, 59 µmol) was added to a 40 mL mixture of 3 M aqueous NaOH and methanol (1:1) and refluxed for 12 h. The solution was acidified with acetic acid and extracted several times with EtOAc until the remaining solution was almost colorless. The combined organic extracts were dried over $MgSO_4$ and concentrated under reduced pressure to provide **3-6** as a purple powder (40 mg, 98% yield). ^{1}H NMR (CD_3OD, 400 MHz): δ 9.99, 9.93, 9.92, 9.91 (4s, 4H); 4.35 (t, J = 7.5 Hz, 4H); 4.03 (q, J = 7.6 Hz, 4H); 3.56, 3.55 (2s, 12H); 3.27 (t, J = 7.5 Hz, 4H); 1.88 (t, J = 7.6 Hz, 6H).

Meso-tetrakis-(4-carboxyphenyl)porphyrin tetramethyl ester (**2-3** or **3-7**) was synthesized following the published procedure.[39]

Cobalt(II) *meso*-tetrakis-(4-carboxyphenyl)porphyrin tetramethyl ester (3-8). The procedure for the synthesis of the structurally related Co(TPP) was applied.[32]

Yield: (384 mg, 90 %). MALDI-MS: *m/z* 903.3 [M^+].

Cobalt(II) *meso*-tetrakis-(4-carboxyphenyl)porphyrin (3-9). Cobalt porphyrin **3-8** (70 mg, 77.5 µmol) was added to a 100 mL mixture of 3 M aqueous NaOH and methanol (1:1) and refluxed for 4 h. The solution was acidified with acetic acid and extracted several times with EtOAc until the remaining solution was almost colorless. The combined organic extracts were dried over $MgSO_4$ and concentrated under reduced

pressure to provide **3-9** as a red powder (63 mg, 96 % yield). MALDI-MS: *m/z* 847.1 [M^+]. MALDI-HRMS calculated for [M^+] $C_{48}H_{28}CoN_4O_8$ 847.1239, found 847.1249.

3.10.3 Reconstitution of Heme Proteins

The procedure for removal of the heme group of Teale was used to generate apoprotein (e.g. myoglobin).[25] The apoprotein was dialyzed against the desired buffer. Generally, about 50 % of the protein survived the harsh treatment. Reconstitution with the selected cofactor was achieved by mixing of the apoprotein with the cofactor (usually 0.9 – 1.0 equivalents).[24]

3.10.4 Reconstitution of Wild Type Myoglobin Monitored by UV-vis Spectroscopy

To ruthenium porphyrin **3-6** (5.8 µmol, 11.6 µM) in PIPES buffer (30 mM, pH 7.0, 0.5 mL) was added one equivalent of apomyoglobin while stirring.

3.10.5 Attempted Cyclopropanation with Myoglobin and Myoglobin Mutants Reconstituted with Ru(CO)MPIX

The procedure of chapter 2.7.6 at pH 7.0 was followed. A typical cyclopropanation reaction was conducted with 1.20 µmol of Ru(CO)Mb (100 µM, 0.14 mol %).

3.10.6 The Influence of Imidazole on the Catalyst Activity of Ru(CO)MPIX

Ruthenium porphyrin **3-6** (129 nmol, 86 µM) was reacted with an equimolar amount of ethyl diazoacetate in 1.5 mL PIPES buffer (100 mM, pH 7.5) in the presence of 3 mM, 30 mM and 300 mM imidazole and the reaction monitored by UV-vis spectroscopy.

The cyclopropanation of styrene with **3-6** was carried out according to chapter 2.7.6 with and without 37.5 mM imidazole and analyzed by GC-FID (chapter 2.7.5).

3.10.7 Cyclopropanation by Cobalt Porphyrins

The cyclopropanation of styrene with **3-9** was carried out according to chapter 2.7.6 under argon atmosphere with and without 37.5 mM imidazole and analyzed by GC-FID (chapter 2.7.5).

3.10.8 Construction of Expression Plasmids of Horse Heart Myoglobin

Myoglobin H64G. The bacterial expression vector pGYM-myoglobin (kind gift from Mauk research group)[40,41] was used as a template for site-directed mutagenesis using overlap extension PCR with oligonucleotides 5'-CGCAGGATTTACATATGGGTCTGTCTGATGGTGAATGG-3', 5'-CGTCTGAAGATCTGAAAAAA**GGT**GGTACCGTTGTGTTAACTGCCCTAGGTGGC-3', 5'-CCTAGGGCAGTTAACACAACGGTACC**ACC**TTTTTTCAGATCTTCAGACGCC-3', 5'-TAAATAACAAAGCTTTTAGTGGTGATGGTGATGGTGACCCGGGCCGCCACCCTGGA

AACCCAGTTCTTTGTACTTAGC-3' and inserted into the *Nco*I and *Hin*dIII sites of pET 28a to give pET 28 Myo H64G.

Myoglobin H93G. The bacterial expression vector pGYM-myoglobin was used as a template for site-directed mutagenesis using overlap extension PCR with oligonucleotides 5'- CGCAGGATTTACATATGGGTCTGTCTGATGGTGAATGG-3', 5'-GCTCAAACCGCTTGCGCAATCG**GGT**GCTACTAAACACAAGATCCCGATC-3', 5'-GGATCTTGTGTTTAGTAGCACCCGATTGCGCAAGCGGTTTGAGCTCAGC-3', 5'-TAAATAACAAAGCTTTTAGTGGTGATGGTGATGGTGACCCGGGCCGCCACCCTGGA AACCCAGTTCTTTGTACTTAGC-3' and inserted into the *Nco*I and *Hin*dIII sites of pET 28a to give pET 28 Myo H93G.

Myoglobin H64G/H93G. The bacterial expression vector pET 28 Myo H93G was used as template for site-directed mutagenesis using overlap extension PCR with oligonucleotides 5'- CGCAGGATTTACATATGGGTCTGTCTGATGGTGAATGG-3', 5'-CGTCTGAAGATCTGAAAAAA**GGT**GGTACCGTTGTGTTAACTGCCCTAGGTGGC-3', 5'-CCTAGGGCAGTTAACACAACGGTACC**ACC**TTTTTTCAGATCTTCAGACGCC-3', 5'-TAAATAACAAAGCTTTTAGTGGTGATGGTGATGGTGACCCGGGCCGCCACCCTGGA AACCCAGTTCTTTGTACTTAGC-3' and inserted into the *Nco*I and *Hin*dIII sites of pET 28a to give pET 28 Myo H64G/H93G.

3.10.11 Construction of Expression Plasmids of Cytochrome P450 (CYP154C1)

CYP154C1 C355A. The bacterial expression vector pet 17-CYP154C1 (kind gift from Waterman research group)[42] was used as a template for site-directed mutagenesis using

overlap extension PCR with oligonucleotides 5'-AGGAGATATACATATGACGACCGGCACCGAAGAAGC-3', 5'- CCTTCGGCCACGGCCCGCACGTG**GCG**CCCGGTGCGGCCCTGTCCCGG-3', 5'-CAGGGCCGCACCGGG**CGC**CACGTGCGGGCCGTGGCCGAAGGAGATGTGC-3' and 5'- CACGTGCGGGCCGTGGCCGAAGG-3' and inserted into the *Bam*HI and *Pml*I sites.

3.10.9 Expression of Myoglobin and Cytochrome P450

The published procedures were followed.[41,42]

3.10.10 Reactions of Myoglobin H64G/H93G Reconstituted with Ru(CO)MPIX Monitored by UV-Vis Spectroscopy

Reaction with EDA. A Ru(CO)Mb 64/93G (purified by size exclusion chromatography) solution (17.5 µM) in PIPES buffer (30 mM, pH 7.5, 530 µL) was reacted with ethyl diazoacetate (2 equivalents) for 2 h.

Reaction with EDA and Styrene. A Ru(CO)MPIX solution (36.6 nmol, 24.4 µM) in PIPES buffer (30 mM, pH 7.5, 1.5 mL) was reacted with EDA (1 equivalent) and styrene (1 equivalent) for one hour. Apomyoglobin 64/93G (1 equivalent) was added to the solution and another equivalent of styrene and EDA and the reaction followed for one hour.

3.10.11 Purification of Myoglobin Reconstituted with 3-6

Generally size-exclusion chromatography was attempted with sephadex G-25 as stationary phase. Cation-exchange chromatography with carboxy methyl cellulose led to precipitation of the protein on the column.

3.10.12 Examination of Non-Specific Binding by ATR-IR

Several mixtures of Ru(CO)MPIX and proteins were measured in phosphate buffer (10 mM; pH 8.0) with the following concentrations: Ru(CO)MPIX (230.7 µM); Ru(CO)MPIX (230.7 µM) – GFP (57.6 µM); Ru(CO)MPIX (420 µM) – myoglobin (210 µM); Ru(CO)MPIX (230 µM) – myoglobin-His (165 µM). The IR-ATR spectra were recorded from 4000 - 400 cm^{-1} using a Bruker Equinox 55 FT-IR spectrometer (Bruker Optics Inc., Billerica, MA) equipped with a liquid nitrogen (LN2) cooled MCT detector (Infrared Associates, Stuart, FL) and a Specac Gateway in-compartment horizontal ATR unit (Specac Inc., Woodstock, GA). Each sample was directly deposited onto a trapezoidal ZnSe (n_D = 2.43 at λ = 5 µm) ATR crystal (MacroOptica, Moscow, Russia) with 6 effective reflection regions (72 × 10 × 6 mm; 45 °).

Purification of expressed proteins: The proteins were purified either by a Ni^{2+}-NTA superflow column or by Ni^{2+}-NTA superflow batch purification according to the Quiagen protocol (The Quiagen expressionist volume 6).

UV-vis spectroscopy: UV-vis absorption spectra were recorded at 25 ℃ u sing a Varian Bio50 UV-vis spectrometer with constant-temperature accessory.

NMR spectroscopy: All NMR spectra (^{1}H, ^{13}C) were recorded on either a Bruker DRX 500 or Mercury VX 400.

GC-FID and GC-MS: An HP 5890 GC equipped with a 6890 series autosampler was used for GC-FID analysis with a DB-5 capillary column. An Agilent GC Chemstation was utilized for peak integration and data analysis. Response factors were calculated for all of the major compounds according to chapter 2.7.5. GC-MS analysis was conducted on an HP 5973 MSD coupled with an HP 5890 GC, using a DB-5 capillary column. Peak identifications were made using the Wiley Registry of Mass Spectral Data (6^{th} edition).

3.11 Literature

(1) Hamachi, I.; Shinkai, S. *European Journal of Organic Chemistry* **1999**, 539-549.

(2) Ozaki, S.; Yang, H. J.; Matsui, T.; Goto, Y.; Watanabe, Y. *Tetrahedron-Asymmetry* **1999**, *10*, 183-192.

(3) Tuynman, A.; Spelberg, J. L.; Kooter, I. M.; Schoemaker, H. E.; Wever, R. *Journal of Biological Chemistry* **2000**, *275*, 3025-3030.

(4) Dexter, A. F.; Lakner, F. J.; Campbell, R. A.; Hager, L. P. *Journal of the American Chemical Society* **1995**, *117*, 6412-6413.

(5) Allain, E. J.; Hager, L. P.; Deng, L.; Jacobsen, E. N. *Journal of the American Chemical Society* **1993**, *115*, 4415-4416.

(6) Zong, Q.; Osmulski, P. A.; Hager, L. P. *Biochemistry* **1995**, *34*, 12420-12425.

(7) Levinger, D. C.; Stevenson, J. A.; Wong, L. K. *Journal of the Chemical Society-Chemical Communications* **1995**, 2305-2306.

(8) Lentz, O.; Qing-Shang, L. I.; Schwaneberg, U.; Lutz-Wahl, S.; Fischer, P.; Schmid, R. D. *Journal of Molecular Catalysis B-Enzymatic* **2001**, *15*, 123-133.

(9) Appel, D.; Lutz-Wahl, S.; Fischer, P.; Schwaneberg, U.; Schmid, R. D. *Journal of Biotechnology* **2001**, *88*, 167-171.

(10) Li, Q. S.; Schwaneberg, U.; Fischer, P.; Schmid, R. D. *Chemistry-a European Journal* **2000**, *6*, 1531-1536.

(11) Schwaneberg, U.; Sprauer, A.; Schmidt-Dannert, C.; Schmid, R. D. *Journal of Chromatography A* **1999**, *848*, 149-159.

(12) Jackson, A. H.; Rao, K. R. N.; Wilkins, M. *Journal of the Chemical Society-Perkin Transactions 1* **1987**, 307-312.

(13) Rillema, D. P.; Nagle, J. K.; Barringer, L. F.; Meyer, T. J. *Journal of the American Chemical Society* **1981**, *103*, 56-62.

(14) Morishima, I.; Shiro, Y.; Nakajima, K. *Biochemistry* **1986**, *25*, 3576-3584.

(15) Ryabova, E. S.; Rydberg, P.; Kolberg, M.; Harbitz, E.; Barra, A. L.; Ryde, U.; Andersson, K. K.; Nordlander, E. *Journal of Inorganic Biochemistry* **2005**, *99*, 852-863.

(16) Cowan, J. A.; Gray, H. B. *Inorganic Chemistry* **1989**, *28*, 2074-2078.

(17) Shiro, Y.; Takeda, M.; Morishima, I. *Journal of the American Chemical Society* **1988**, *110*, 4030-4035.

(18) Paulson, D. R.; Addison, A. W.; Dolphin, D.; James, B. R. *Journal of Biological Chemistry* **1979**, *254*, 7002-7006.

(19) Srivastava, T. S. *Biochimica Et Biophysica Acta* **1977**, *491*, 599-604.

(20) Horrocks, W. D.; Venteicher, R. F.; Spilburg, C. A.; Vallee, B. L. *Biochemical and Biophysical Research Communications* **1975**, *64*, 317-322.

(21) Muir, H. M.; Neuberger, A. *Biochemical Journal* **1949**, *45*, 163-170.

(22) Reboucas, J. S.; James, B. R. *Tetrahedron Letters* **2006**, *47*, 5119-5122.

(23) Reboucas, J. S.; Cheu, E. L. S.; Ware, C. J.; James, B. R.; Skov, K. A. *Inorganic Chemistry* **2008**, *47*, 7894-7907.

(24) Yonetani, T. *Journal of Biological Chemistry* **1967**, *242*, 5008-&.

(25) Teale, F. W. J. *Biochimica Et Biophysica Acta* **1959**, *35*, 543-543.

(26) Breslow, E.; Koehler, R. *Journal of Biological Chemistry* **1965**, *240*, 2266-&.

(27) Breslow, E.; Beychok, S.; Hardman, K. D.; Gurd, F. R. N. *Journal of Biological Chemistry* **1965**, *240*, 304-&.

(28) Lee, V. W. S.; Chen, Y. L.; Konermann, L. *Analytical Chemistry* **1999**, *71*, 4154-4159.

(29) Niimi, T.; Uchida, T.; Irie, R.; Katsuki, T. *Advanced Synthesis & Catalysis* **2001**, *343*, 79-88.

(30) Csuk, R.; Schabel, M. J.; vonScholz, Y. *Tetrahedron-Asymmetry* **1996**, *7*, 3505-3512.

(31) Huang, L. Y.; Chen, Y.; Gao, G. Y.; Zhang, X. P. *Journal of Organic Chemistry* **2003**, *68*, 8179-8184.

(32) Adler, A. D.; Longo, F. R.; Varadi, V. *Inorganic Syntheses*, 1976; Vol. 16.

(33) Roach, M. P.; Ozaki, S.; Watanabe, Y. *Biochemistry* **2000**, *39*, 1446-1454.

(34) Barrick, D. *Biochemistry* **1994**, *33*, 6546-6554.

(35) Liong, E. C.; Dou, Y.; Scott, E. E.; Olson, J. S.; Phillips, G. N. *Journal of Biological Chemistry* **2001**, *276*, 9093-9100.

(36) Goldberg, M. E.; Chaffotte, A. F. *Protein Science* **2005**, *14*, 2781-2792.

(37) Komatsu, T.; Matsukawa, Y.; Tsuchida, E. *Bioconjugate Chemistry* **2002**, *13*, 397-402.

(38) Nimri, S.; Keinan, E. *Journal of the American Chemical Society* **1999**, *121*, 8978-8982.

(39) Dattagup.N; Bardos, T. J. *Journal of Heterocyclic Chemistry* **1966**, *3*, 495-&.

(40) Lloyd, E.; Mauk, A. G. *Febs Letters* **1994**, *340*, 281-286.

(41) Guillemette, J. G.; Matsushimahibiya, Y.; Atkinson, T.; Smith, M. *Protein Engineering* **1991**, *4*, 585-592.

(42) Podust, L. M.; Kim, Y.; Arase, M.; Neely, B. A.; Beck, B. J.; Bach, H.; Sherman, D. H.; Lamb, D. C.; Kelly, S. L.; Waterman, M. R. *Journal of Biological Chemistry* **2003**, *278*, 12214-12221.

Chapter IV

DESIGN OF SEMI-SYNTHETIC PROTEINS FOR THE STUDY OF LABILE POOLS OF COPPER, IRON AND ZINC IN INTRACELLULAR COMPARTMENTS

4.1 Introduction and Concept

Green fluorescent protein (GFP) has revolutionized cellular imaging and cell biology (see chapter 1.2.5). It offers a non-invasive method for imaging virtually any protein of interest within live cells based on fluorescence microscopy. The fluorophore of GFP is buried inside of a cylindrical β-barrel, which renders the fluorescence emission mostly insensitive towards environmental changes, in particular solvent polarity. At the same time, this insensitivity poses significant challenges for the design of GFP based biosensors that will respond through the fluorescence intensity shift towards analyte binding. The binding of metals to the surface of the protective protein framework results in only small changes of the fluorescence emission due to the strong distance dependence of the energy transfer efficiency (see chapter 1.2.7).[1] That is why GFP cannot be converted to an efficient fluorescent metal sensor for *in vivo* applications. In contrast small molecule fluorescent probes can be used that exhibit a change of fluorescence upon metal binding, which subsequently allows the determination of the presence or concentration of metals; however, small molecule fluorescent probes usually have no anticipated specificity for any proteins or intracellular locations. A classic

example for their “mind of their own” is rhodamine B, which tends to accumulate in the mitochondria.[2] On the other hand, proteins such as GFP can be directed to intracellular compartments by signal peptides (see chapter 1.2.6) or by fusion to a protein with specific subcellular localization. One of the biggest challenges in metal sensing based on small molecule fluorescent dyes is that open d-shell cations from transition metals, such as copper and iron, quench the fluorophore fluorescence efficiently. Colorimetric and complexometric methods for the qualitative and quantitative determination of transition metals were widely used before the emergence of modern analytical instrumental techniques such as atomic absorption spectroscopy (AAS) or inductively coupled plasma mass spectrometry (ICPMS).[3,4] The formation of colored chelator-metal charge transfer complexes, with spectral properties that are highly dependent on the bound metal, allows the analytical determination of metal ions by colorimetry. Therefore, metal chelators were developed that form colored charge transfer complexes for the metals copper and iron, both of which are typically strong fluorescence quenchers (Figure 4-1). Unfortunately, these charge transfer complexes cannot be used for live cell imaging applications because they do not emit any light.

Nevertheless, the combination of colored charge transfer complexes with a signal peptide carrying fluorescent protein or the combination of a signal peptide carrying protein with a small molecule fluorescent chelator by generating semi-synthetic proteins may take advantage of beneficial properties of each species while bypassing their downsides. Our design of the semi-synthetic protein for the study of labile metal pools is based on a genetically encoded fluorescent protein that is modified at the N- and C-terminus (Figure 4-2). A signal peptide directs the protein and therefore the attached metal sensor to intracellular compartments of interest; thus, the problem of non-existing

or unwanted specificity of small molecule fluorescent sensors can be circumvented. A number of organelles, including nucleus, mitochondria and the Golgi apparatus (and the *trans* Golgi network) are likely locations of labile metal pools and therefore were chosen as target locations (Figure 4-3).[5-7] The fluorescent protein is expressed as a fusion protein; the fusion domain allows labeling of the protein. Labeling of the protein with the metal chelator by AGT and by *trans*-splicing with split-inteins was chosen (see chapter 1.2.8) because both labeling methods are fairly general with high substrate tolerance and are also suitable for *in vivo* studies.[8-11] The protein has to be labeled with a metal chelator that either exhibits a suitable change of fluorescence or color upon binding of the desired metal. This change results in the formation of a Förster resonance energy transfer (FRET) pair in which the fluorescent protein acts as the donor fluorophore and the colored species as acceptor (in case of a fluorescent metal chelator complex this relationship can also be the other way around). The FRET pair can then be analyzed by fluorescence emission of the donor-acceptor pair or fluorescence lifetime of the donor (the fluorescence lifetime of the acceptor does not change). Thus, the local metal concentrations can be determined while circumventing the notorious fluorescence quenching of copper and iron (for mechanism see chapter 4.1.1). In case of sensing Zn(II) (which does not quench the fluorescence efficiently), the design can be slightly changed if the metal chelator exhibits fluorescence upon metal binding; the fluorescent protein domain can be omitted since another fluorophore is not required. The fusion protein can be expressed inside of the target mammalian cells after transfection with generated plasmids encoding the protein. Incubation of the cells with the precursor metal chelator (contains moiety that allows attachment to protein) should lead to the *in vivo* labeling of the protein, thus forming the metal sensor.

Many metal chelators form complexes with 2:1 or even 3:1 ligand-metal stochiometry (Figure 4-1). For sensing applications a 1:1 binding stochiometry is preferable because a variety of lewis bases in the cytoplasm (e.g. amines and thiols) can coordinate to the metal center. The resulting heteroleptic metal complexes would distort live cell measurements. Utilizing semi-synthetic proteins labeled with 2:1 or 3:1 binding ligands would always lead to potential coordination sites for intracellular Lewis bases because the steric hindrance of the protein domain does not allow the required movement of two or three semi-synthetic proteins around one metal cation to form a complex. Hence, 1:1 binding chelators are required for intracellular measurements. A crucial property of the metal chelator for the goal of *in vivo* cell imaging of the labile metal pools of copper, iron and zinc is the appropriate binding affinity for the desired metal. Too weak binding would lead to a low signal while too tight binding would simply strip metals from metalloproteins instead of binding metals of the targeted labile metal pool. A binding affinity in the µM range is suitable because it is a compromise between the two scenarios.

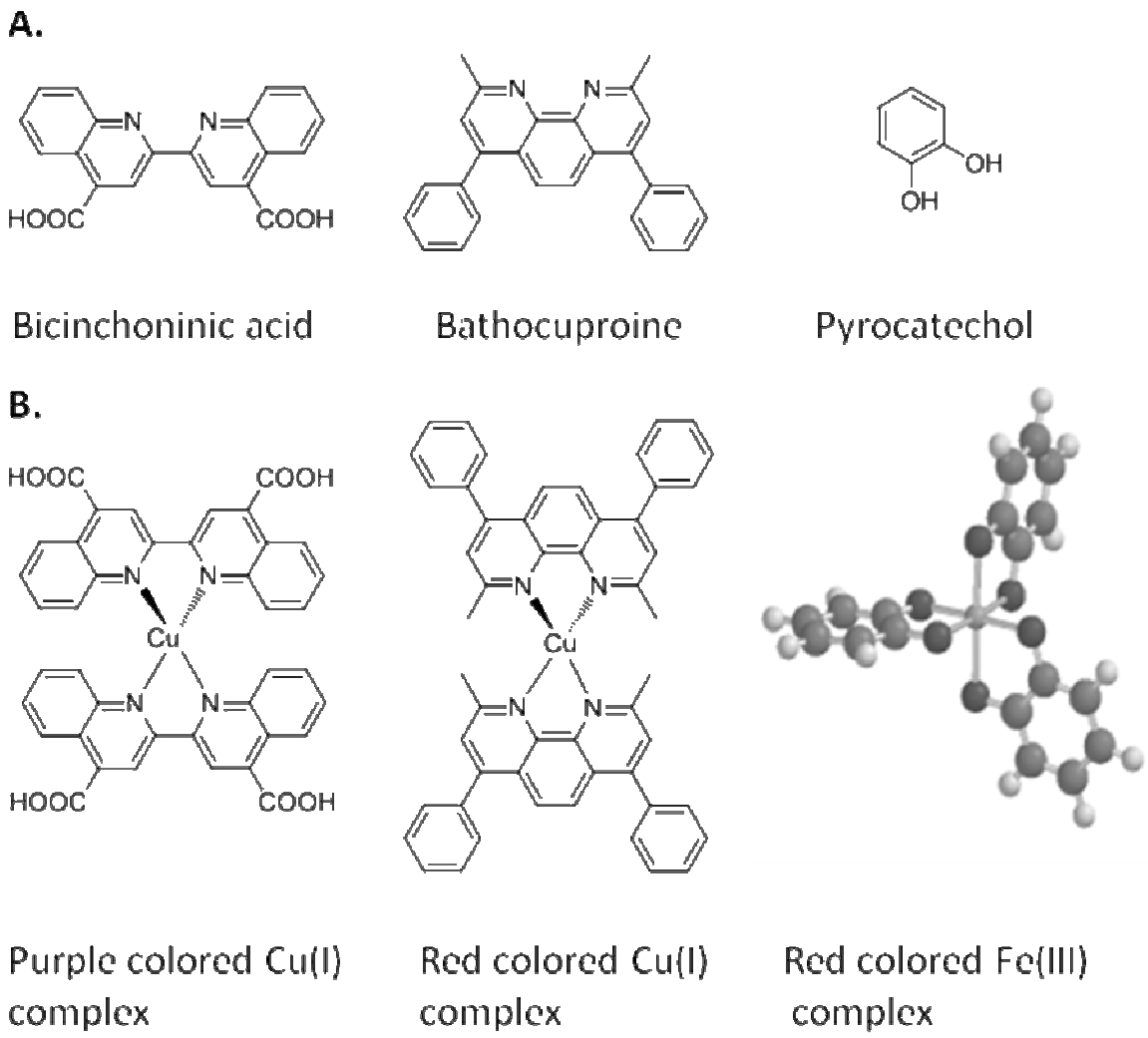


Figure 4-1. Strongly colored metal charge transfer complexes of copper and iron. **A.** free ligands. **B.** metal complexes, which almost exclusively exhibit 2:1 or 3:1 binding behavior (ligand : metal).

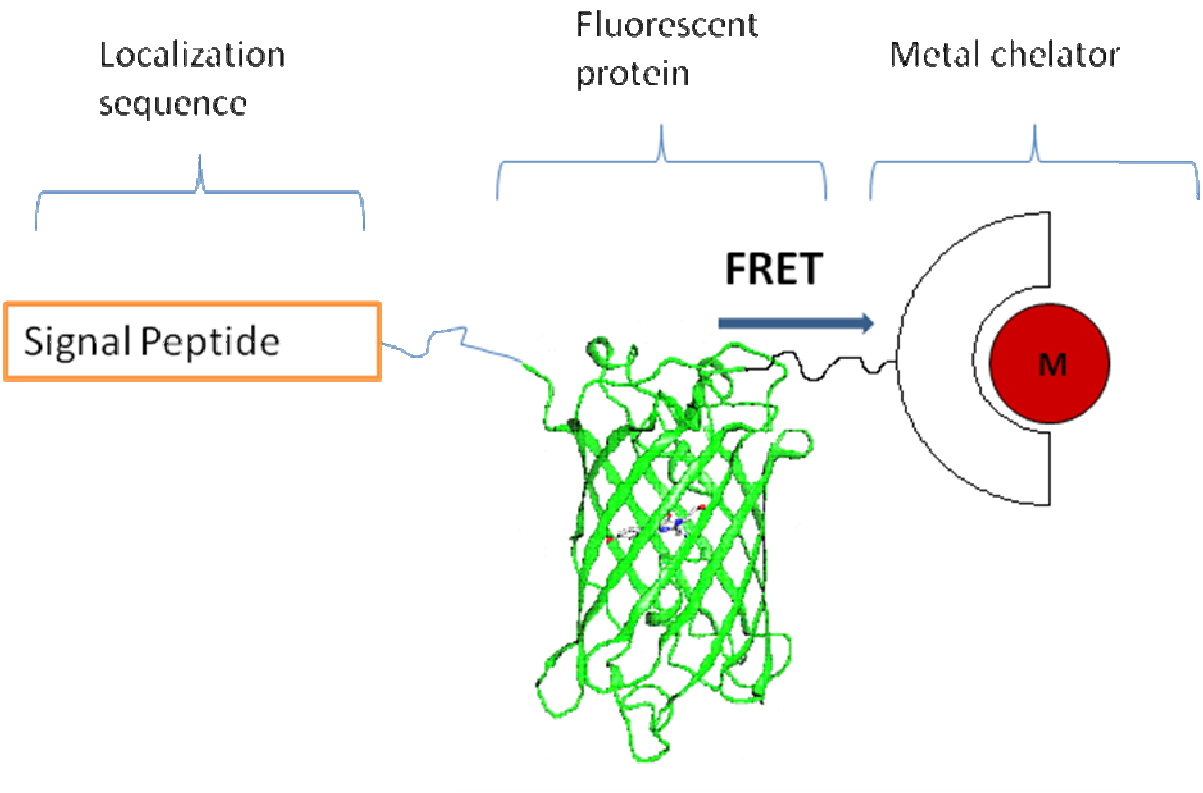


Figure 4-2. Design concept of a genetically encoded, semi-synthetic metal sensor. A fusion protein containing a signal peptide and fluorescent protein domain is expressed in mammalian cells. This fusion protein is labeled with a metal chelator that exhibits color or fluorescence upon metal chelation. Thus, binding of a metal-ion results in the formation of a FRET pair. Analysis of the FRET pair (fluorescence emission and fluorescence lifetime) allows the determination of the local metal concentration of the labile metal pool.

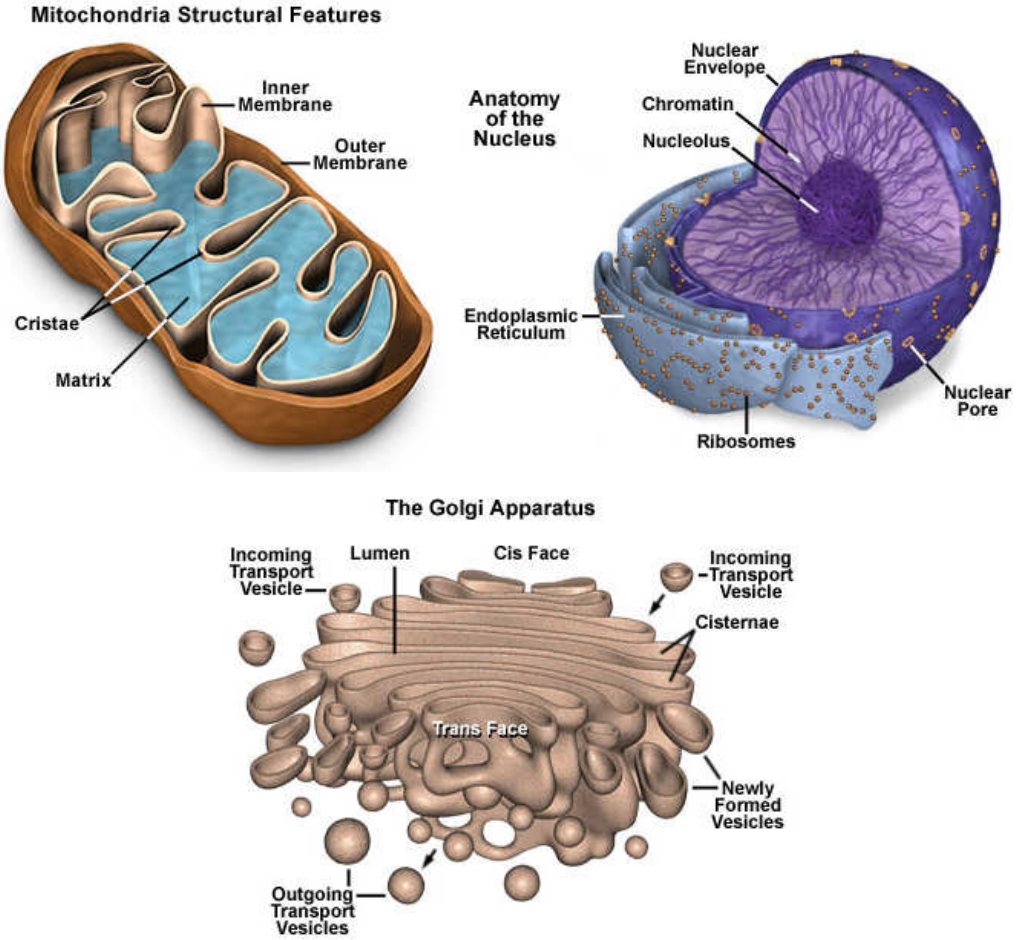


Figure 4-3. Target organelles.

4.1.1 Mechanism of Fluorescence Quenching by Charge Transfer States

The design of the metal sensor is based on FRET of the donor fluorophore to the colored species that forms upon metal binding. Upon absorption of photons with suitable energy the donor fluorophore is excited from the ground state S_0 into the excited state S_1 (Figure 4-4). In the absence of a metal ligand charge transfer (MLCT) state the excited state decays to the ground state either by radiative deactivation under emission of a photon or by non-radiative deactivation through vibrational relaxation. Hence, the quantum yield (Q) and fluorescence lifetime (T_F) of the donor are determined by the rate of radiative (k_r) and non-radiative (k_{nr}) deactivation (equation 4-1).[12] In case of the presence of a low-lying MLCT state there is an additional pathway for the non-radiative deactivation of the excited state. Upon energy transfer (for requirements of energy transfer see chapter 1.2.7) of the excited state of the donor to the MLCT state of the acceptor the excited acceptor state is subsequently deactivated by non-radiative deactivation to the ground state of the acceptor. Thus, the quantum yield and fluorescence lifetime of the donor is dependent on the FRET efficiency. The more efficient the energy transfer to the colored acceptor the shorter is the fluorescence lifetime and the lower is the quantum yield (equation 4-2 and 4-3).

$$Q = \frac{k_r}{k_r + k_{nr}} \qquad \text{(equation 4-1)}$$

$$Q' = \frac{k_r}{k_r + k_{nr} + k_{FRET}} \qquad \text{(equation 4-2)}$$

$$\tau_F' = \frac{1}{k_r + k_{nr} + k_{FRET}} \qquad \text{(equation 4-3)}$$

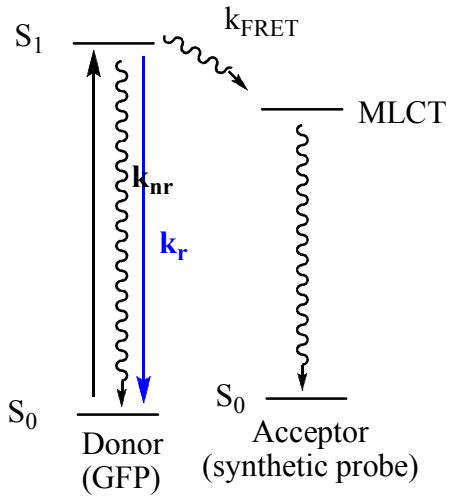


Figure 4-4. Jablonski diagram of fluorescence quenching by FRET. Upon excitation to the excited state S_1 radiative or non-radiative deactivation to the ground state S_0 occurs. In case of the presence of a FRET pair, an additional non-radiative pathway is available. This pathway allows the transfer of energy from the donor (FP) to the acceptor (synthetic probe), thus competing with the deactivation pathways of S_1 to S_0 of the donor. This results in the quenching of the donor fluorescence emission. The efficiencies of the quenching depends on the individual rate constants of the pathways (k_r = rate constant of radiative deactivation; k_{nr} = rate constant of non-radiative deactivation; k_{FRET} = rate constant of donor - acceptor energy transfer).

4.2 Design and Generation of Fusion Proteins

Plasmids are circular extra-chromosomal DNA molecules that are self-replicating and can be transferred from one organism to another.[13] They can be found in both pro- and eukaryotic cells and are essential in genetic engineering as vectors of genes. Therefore plasmids encoding the desired fusion proteins were generated utilizing methods and techniques of molecular cloning, such as transformation, restriction digestions, PCR (polymerase chain reaction) techniques and ligations. The generated plasmids usually contain a localization sequence and a fusion domain enabling protein labeling. The plasmids are designed for the expression of the genes in either eukaryotic or prokaryotic hosts. All of the following depicted plasmids are selected "final" plasmids.

4.2.1 Plasmids for Eukaryotic Expression Targeting Mitochondria Based on Labeling by AGT

The generated plasmids contain a CMV (cytomegalovirus) promoter for efficient eukaryotic expression (Figure 4-5). Plasmid **A** encodes a fusion protein with the genes for AcGFP (a monomeric GFP mutant) and AGT (commercial name SNAP-tag). The genes are connected by a linker encoding 10 amino acids (ALAAADIKLT). The mitochondrial targeting sequence (MTS) is located at the N-terminus of the fusion protein and cannot be placed at the C-terminus. The FRET efficiency is dependent on the overlap integral between donor and acceptor (see chapter 4.3). In order to achieve better spectral overlap with certain compounds, such as bicinchoninic acid derivatives, plasmid **B** was generated that contains a cerulean gene (mutant of cyan fluorescent protein) in place of AcGFP. Theoretically, the order of AcGFP and AGT can be reversed;

however, no obvious advantage of this strategy could be inferred from the crystal structures of the proteins. Successful localization in the mitochondria upon transfection of mammalian cells with plasmid **A** (MTS-AcGFP-AGT) was observed (Figure 4-6).

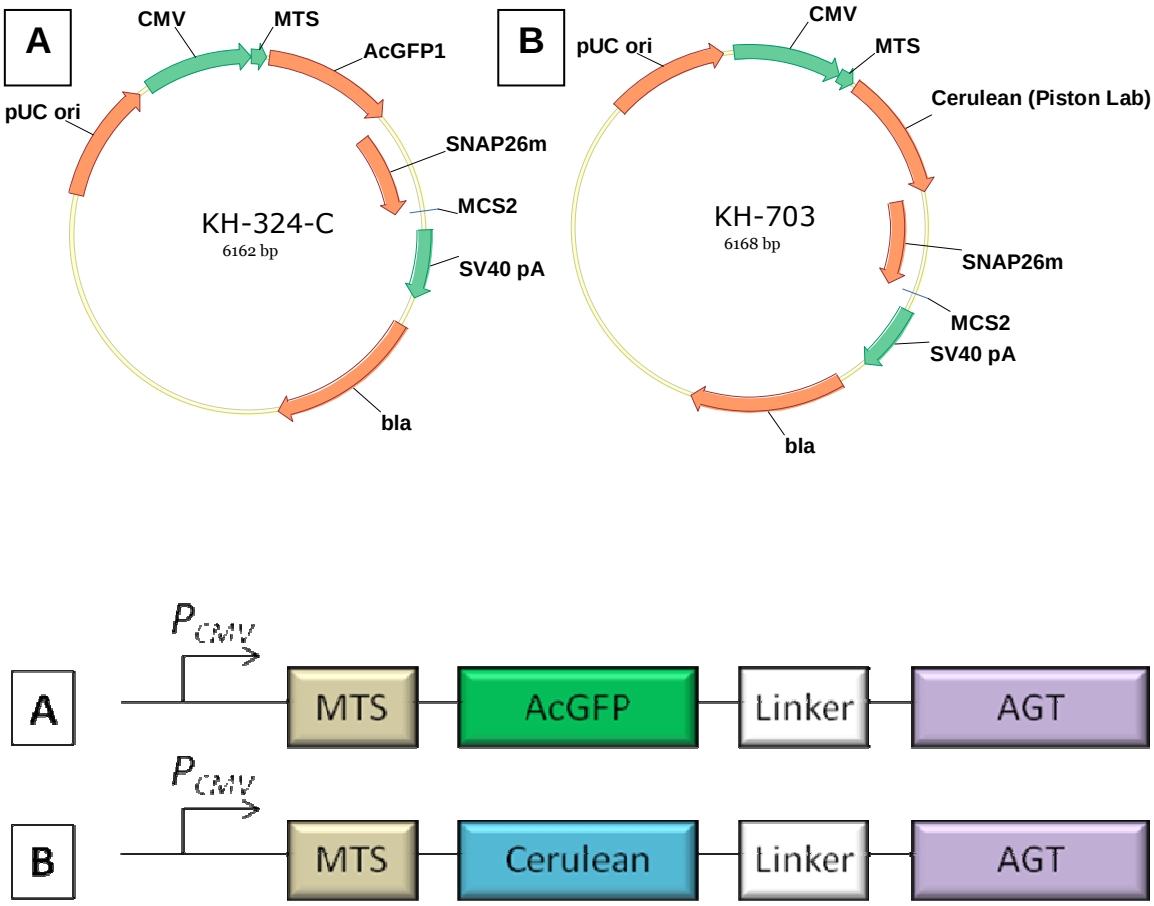


Figure 4-5. Circular and linear representation of plasmids for eukaryotic expression targeting mitochondria.

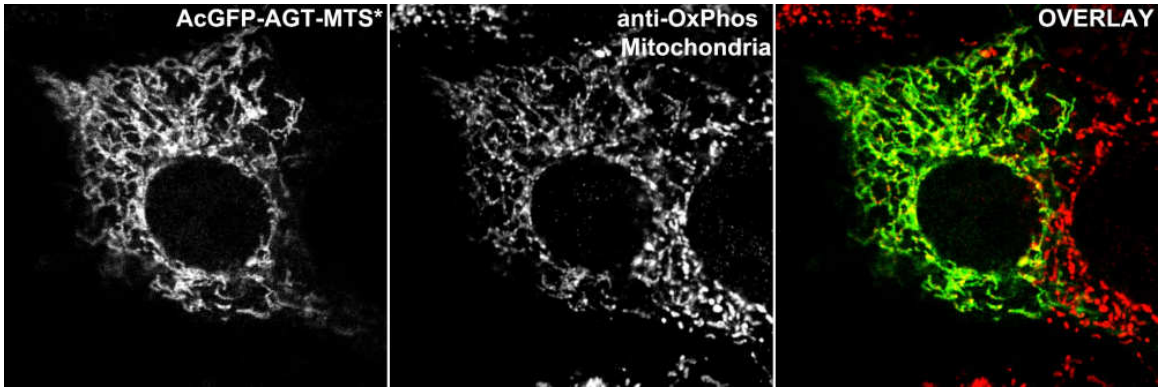


Figure 4-6. Light micrographs of 3T3 cells transfected with plasmid **A** expressing MTS-AcGFP-AGT. Localization in the mitochondria was observed (left panel) and confirmed by staining with mitochondria specific antibodies (center panel) and overlay of the resulting images (right panel).

4.2.2 Plasmids for Eukaryotic Expression Targeting the Nucleus Based on Labeling by AGT

Plasmid **E** is similar to **A** because it encodes AcGFP-AGT with the same linker but has some key differences (Figure 4-7). The leader sequence (sequence between CMV promoter and first start codon) is a Kozak consensus sequence for enhanced expression. The nuclear localization signal is at the C-terminus but may have also been placed at the N-terminus.[14] Localization in the nucleus was confirmed by transfection of mammalian cells with plasmid **E** and subsequent microscopy (Figure 4-8).

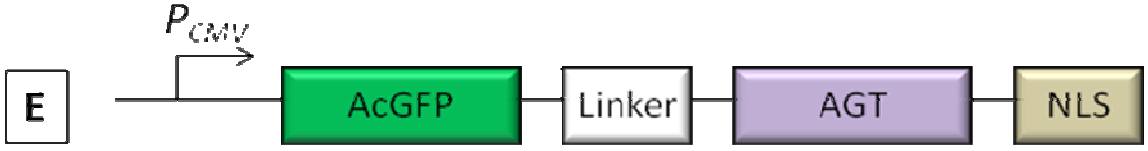


Figure 4-7. Linear representation of plasmids for eukaryotic expression targeting the nucleus.

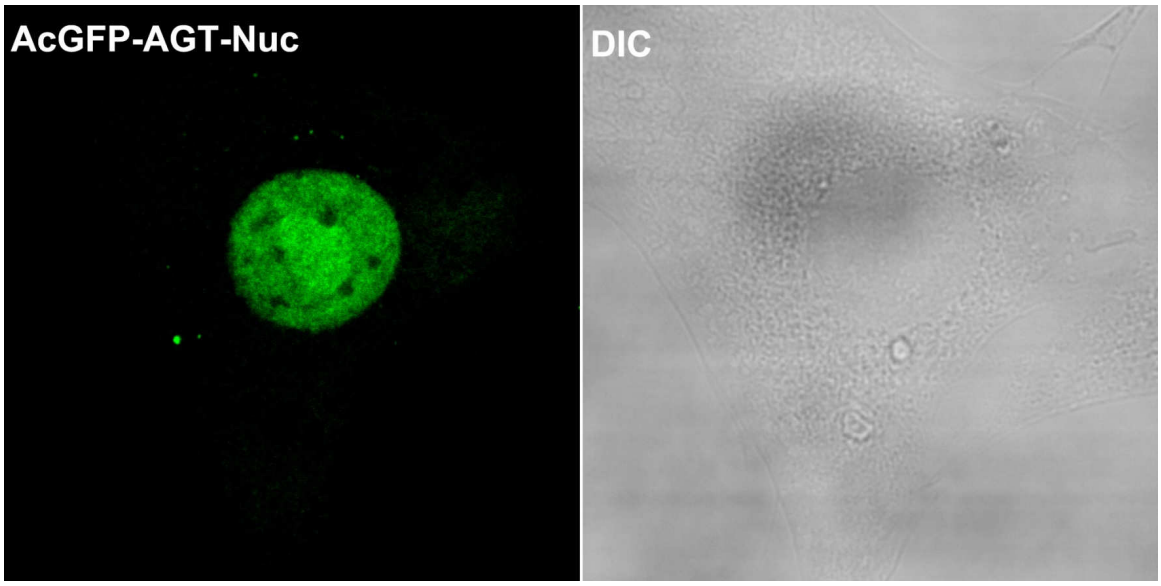


Figure 4-8. Light micrographs of 3T3 cells transfected with plasmid **E** expressing AcGFP-AGT-NLS. Localization in the mitochondria was observed.

4.2.3 Plasmids for Eukaryotic Expression Targeting the *Trans*-Golgi Network

Proteins can be directed to the *trans*-Golgi network (TGN) by expression as fusion protein with TGN38, a membrane protein of the TGN.[15] The advantage of utilizing TGN38 is that the N-terminal domain of this membrane protein is located in the lumen.[16] Therefore the metal ions of the intracellular compartments instead of the cytosol would be analyzed; however, utilizing this strategy would require expressing a triple fusion protein. TGN38 is about the size of the whole AcGFP-AGT domain and could therefore lead to perturbations of the folding process, instability of the fusion protein, etc. Moreover, TGN38 is heavily glycosylated and the impact of this feature on the fusion protein and analyte sensing is hard to determine.[17] Due to the complexity of biological systems a simpler approach to direct the protein to the TGN is preferable. Deletion analysis has revealed that an 11 amino acid segment of TGN38 is sufficient for TGN

localization.[16] Hence, several plasmids encoding AcGFP were generated that contained short TGN localization sequences (TGN-LS) at the C-terminus.

The 11-amino acid encoding sequence was inserted into pAcGFP1-C1 with either an immediate stop codon or additional 20 amino acids before termination of translation with a 5 amino acid linker (SGLRS) (Figure 4-9). Unfortunately, upon transfection of mammalian cells (HeLa cells) no specific localization of the expressed AcGFP in the TGN was observed (Figure 4-10).[18]

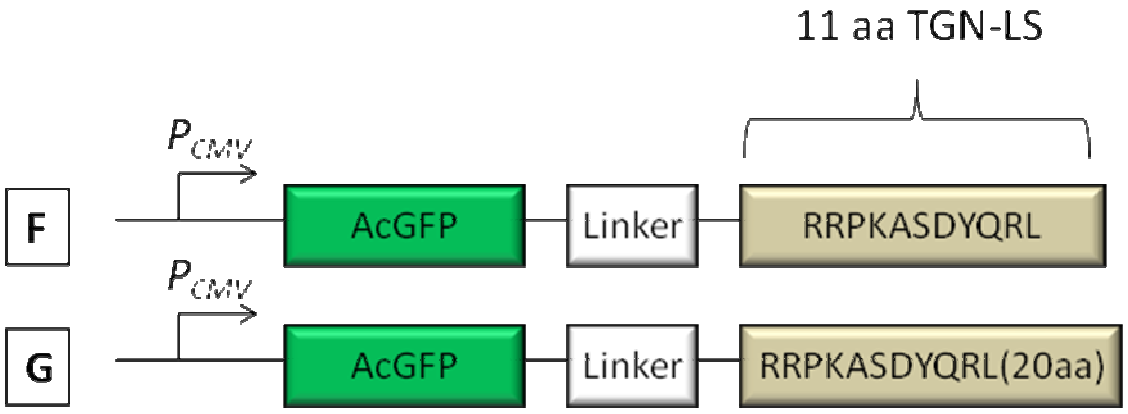


Figure 4-9. Linear representation of plasmids encoding AcGFP with 11 amino acid TGN localization sequences.

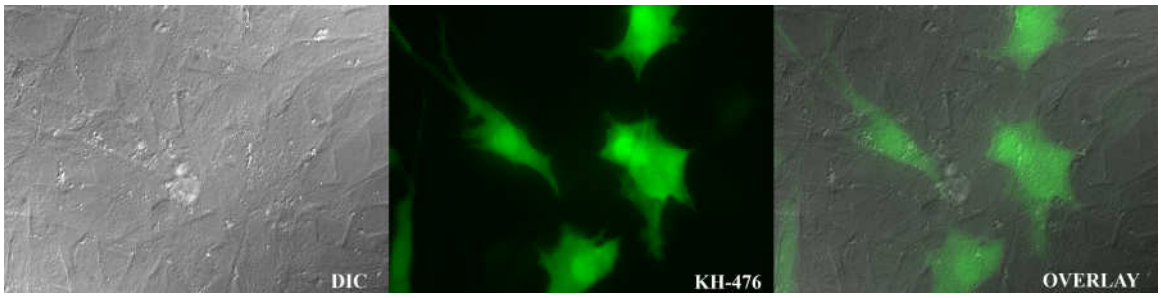


Figure 4-10. Light micrographs of 3T3 cells transfected with **F** expressing AcGFP. Localization in the TGN was not observed.

Due to the failed localization of AcGFP in the TGN the design was modified (Figure 4-11). The first approach was to insert more than 11 amino acids of the TGN38 "tail" assuming that interactions of the localization sequence with the original neighboring amino acids could play a role with regard to the recognition of the signal (**I**). The second approach focused on the generation of repeat localization sequences to increase the probability of being recognized (**H**). This is a common strategy for other localization signals (e.g. nuclear localization sequence).[14] Unfortunately, these designs also failed to localize AcGFP in the TGN. Pap and coworkers have reported on the successful targeting of fluorophores to the TGN utilizing short peptides by the truncated TGN38 localization sequence (TGN38-LS) SDYQRL.[19] The net charge of the localization signal plays an important role in targeting of intracellular compartments. In case of TGN38-LS and adjacent amino acids should be around 0. The basic amino acids, such as arginine in the 11-amino acid localization sequence result in an increase of the net charge. Therefore a SDYQRL double repeat was inserted at the C-terminus of AcGFP1-C1 (**J**). Unfortunately, upon transfection of 3T3 mammalian cells with plasmid **J** localization of AcGFP at the *trans*-Golgi network was not observed either.

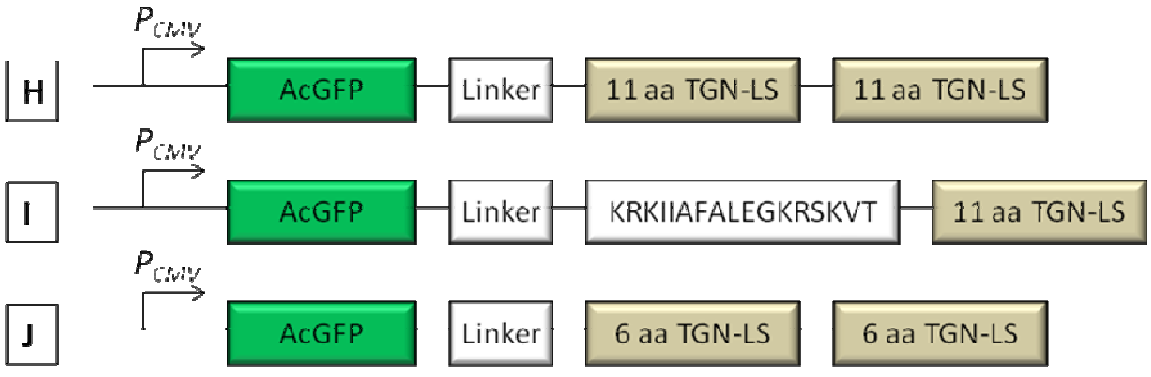


Figure 4-11. Linear representation of plasmid encoding AcGFP with TGN localization sequences.

As a consequence of the failures to direct AcGFP to the TGN by short localization signals, plasmids had to be generated that encode AcGFP and the membrane protein TGN38 as fusion protein or even triple fusion protein (in combination with AGT or intein^{C}) (Figure 4-12 and Figure 4-13). This strategy led to the localization of expressed proteins in the TGN (Figure 4-14). Plasmids **K** and **L** only differ from each other with regard to the fluorescent protein (AcGFP in case of **K** and cerulean in case of **L**). They both have a CMV promoter and a chimeric intron and encode intein^{C}-FP-TGN38. It has been shown that the combination of CMV promoter and chimeric introns usually lead to significantly higher levels of expression.[20] The split-intein^{C} (from *Ssp* DNA B intein) design is based on an investigation of suitable split-intein sites, which revealed that an intein^{N} fragment of only 11 amino acids and its 143 amino acid intein^{C} counterpart were capable of protein *trans*-splicing *in vivo*.[9] Mootz et al. adapted this strategy in order to ligate synthetic peptides to the N-terminus of a recombinant protein.[8] This design constitutes an excellent starting point for attaching metal chelators to the protein, thus generating semi-synthetic proteins for metal sensing. Since the intein^{C} fragment is required to be N-terminal only signal sequences that are functional at the C-terminus, which is a minority of the localization signals, can be utilized.

Plasmid **M** contains a CMV promoter with Kozak consensus sequence and the genes for AcGFP-AGT-TGN38. The AGT based labeling method allows more flexibility with regard to placing signal peptides at the termini.

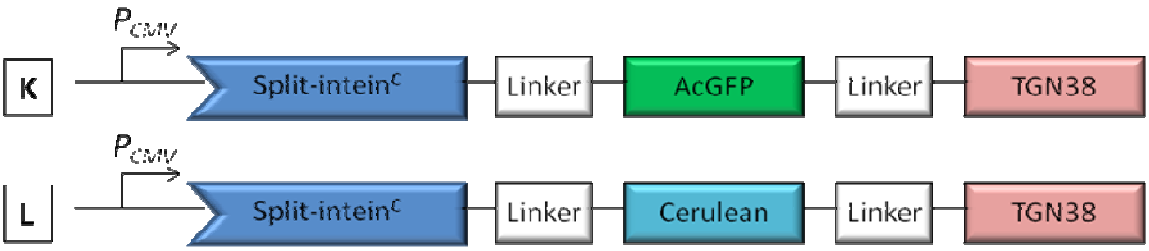


Figure 4-12. Linear representation of split-intein based plasmids targeting TGN.

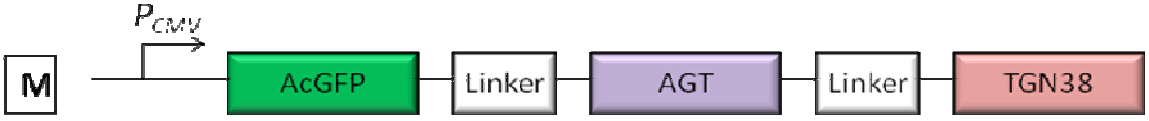


Figure 4-13. Linear representation of AGT based plasmids targeting TGN.

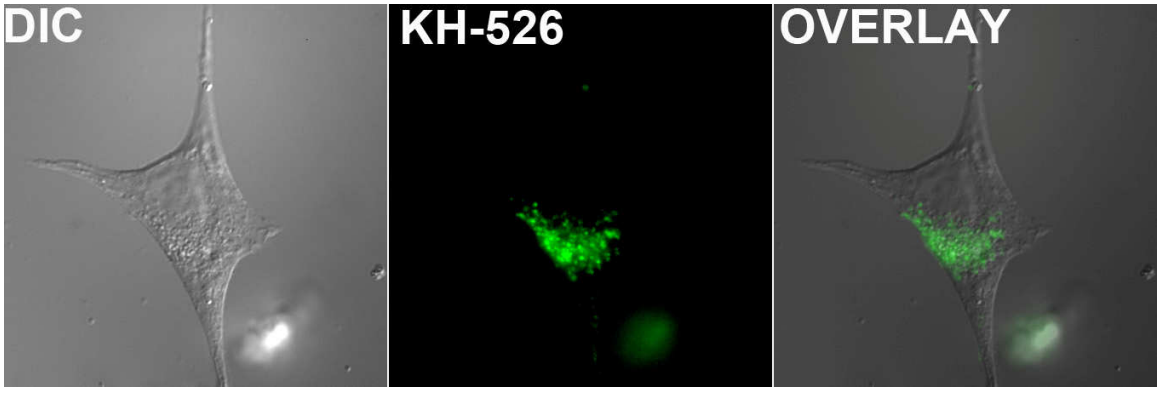


Figure 4-14. Light micrographs of 3T3 cells transfected with plasmids encoding AcGFP-TGN38. Localization in the TGN was observed.

4.2.4 Plasmids for Prokaryotic Expression

The genes of several generated eukaryotic expression plasmids (section 4.2.1 – 4.2.3) were transferred to prokaryotic expression vectors (Novagen pET vector series)

containing T7 promoters for efficient overexpression in bacteria yielding plasmids **N** to **T**. Overexpression in E. coli is necessary in order to obtain suitable amounts of protein for *in vitro* studies. The His-tag that is usually encoded to enable facile purification of the expressed protein might interfere with the in vitro metal sensing experiments due to its ability to strongly bind divalent cations such as nickel, zinc, etc. Therefore plasmid **N** does not encode a His-tag while plasmids **O** to **T** encode either an N-or C-terminal His-tag. In addition plasmid **O** and **P** encode Thrombin protease sites that allow the cleavage of the His-tag. Plasmids **N** to **P** encode AGT (Figure 4-15) while plasmids **Q** to **T** contain split-intein intein^{C} as labeling method (Figure 4-16 and Figure 4-17).

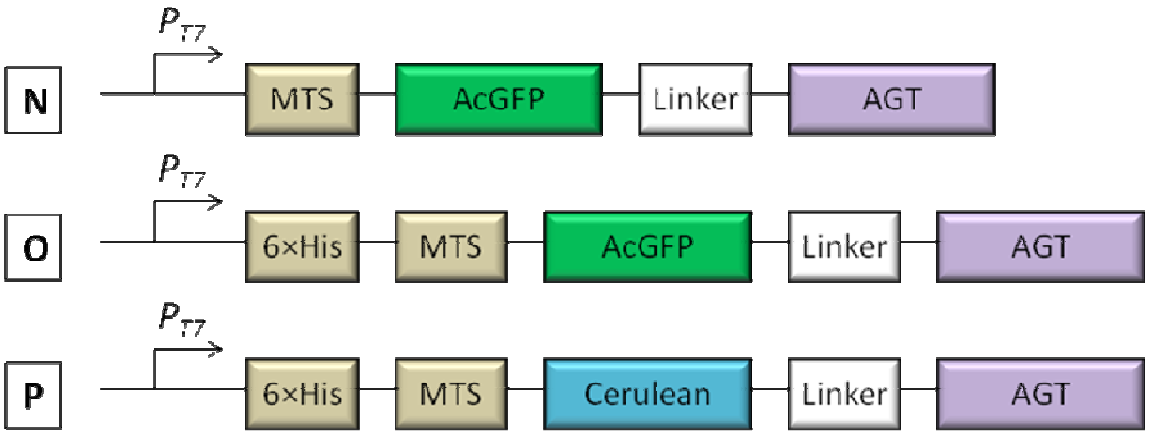


Figure 4-15. Linear representation of bacterial expression plasmids based on labeling by AGT.

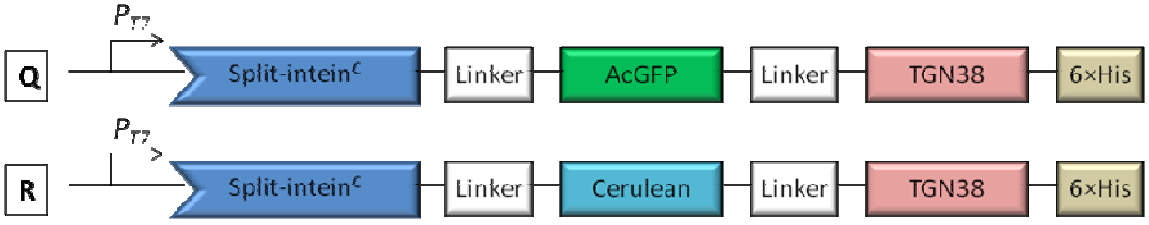


Figure 4-16. Linear representation of bacterial expression plasmids based on labeling by split-inteins.

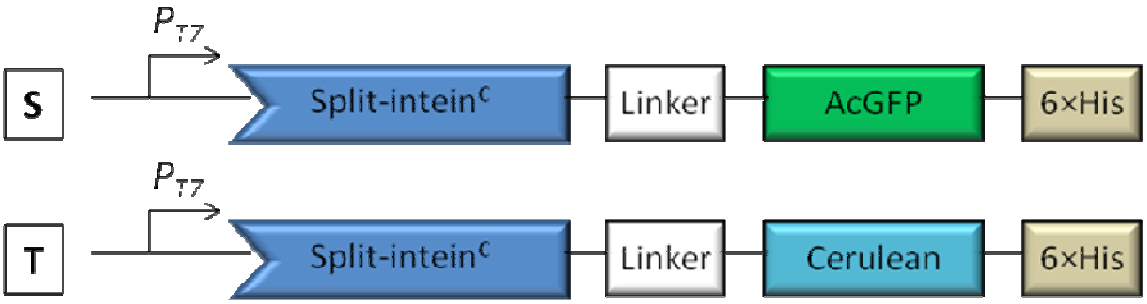


Figure 4-17. Linear representation of bacterial expression plasmids based on labeling by split-inteins.

4.3 *In Vitro* Studies of Labeling Expressed Fusion Proteins with Model Compounds

4.3.1 Labeling of AGT Based Fusion Proteins

4.3.1.1 Expression and Isolation of AGT Based Fusion Proteins

E. coli bacteria were transformed with plasmids for bacterial expression encoding MTS-AcGFP-AGT (plasmids **N** and **O**). The formation of the fusion proteins (MW ≈ 54 kDa) was observed after induction of expression followed by lysis of the bacterial cells (Figure 4-18). However, the desired protein was only present in the insoluble fractions of the cell lysate. Expression of soluble protein was achieved by modification of the standard overexpression conditions to significantly lower temperatures and longer expression times (see experimental section). The desired pure protein was obtained by Ni-NTA (Quiagen) based His-tag purification (Figure 4-19). Expression of the untagged protein was not necessary anymore because the potentially problematic His-tag could be easily removed by proteolysis with thrombine (Figure 4-20). Complete cleavage was

observed at all dilution levels of thrombin (even with 200 fold dilution of the recommended amount by the manufacturer).

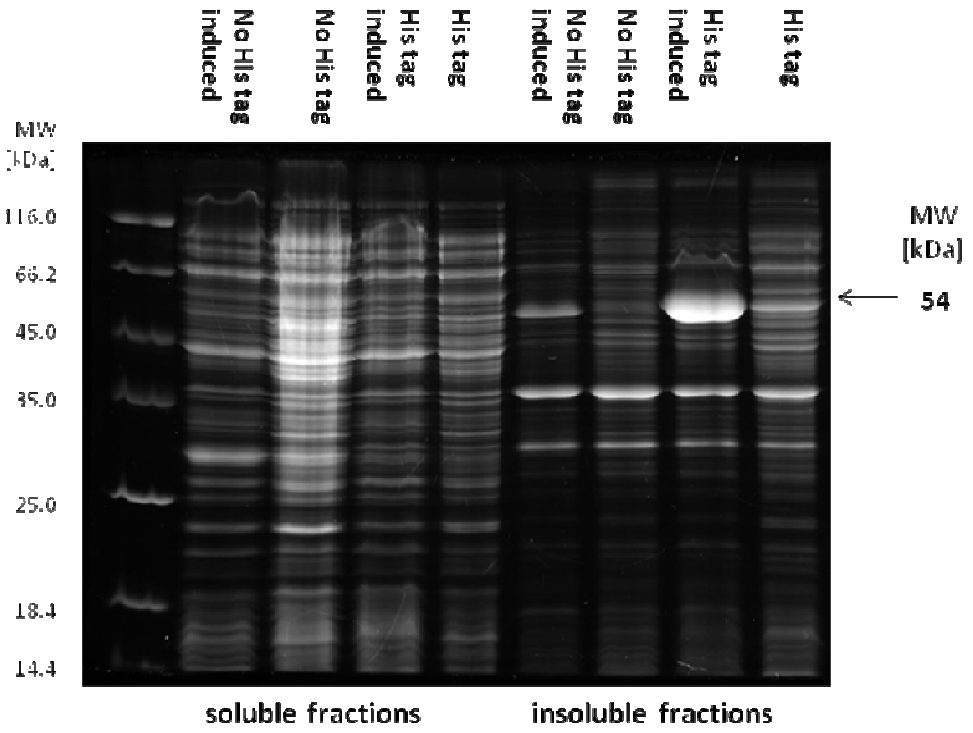


Figure 4-18. SDS-PAGE gel of expression of AGT based fusion proteins (cell lysates). The expressed fusion proteins only appeared in the insoluble fractions.

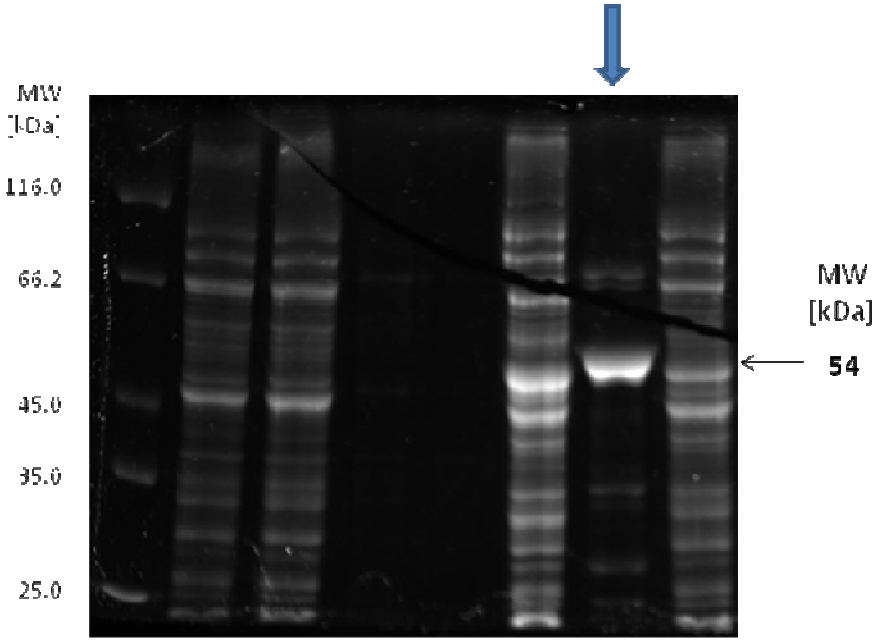


Figure 4-19. SDS-PAGE gel of purified protein MTS-AcGFP-AGT.

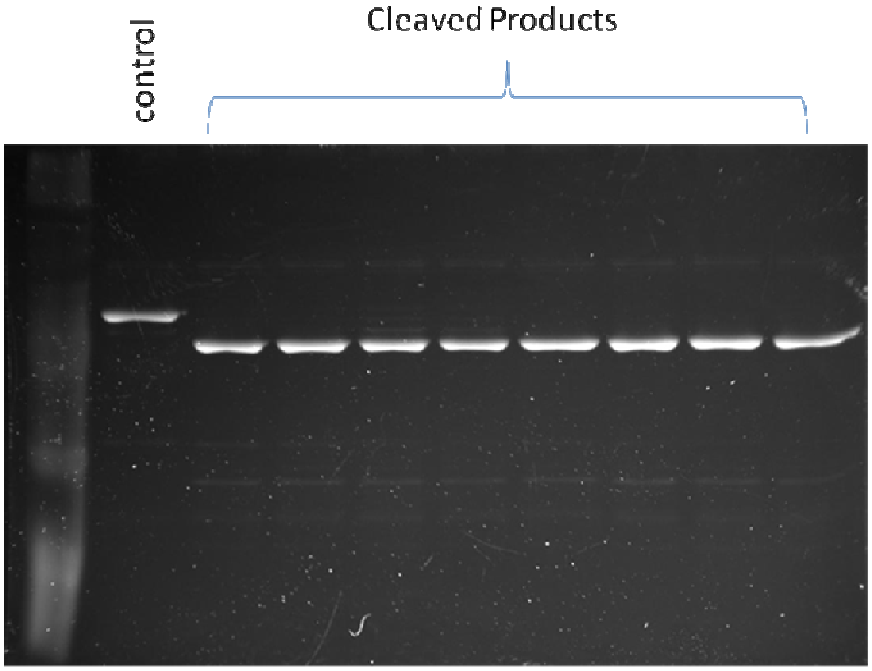
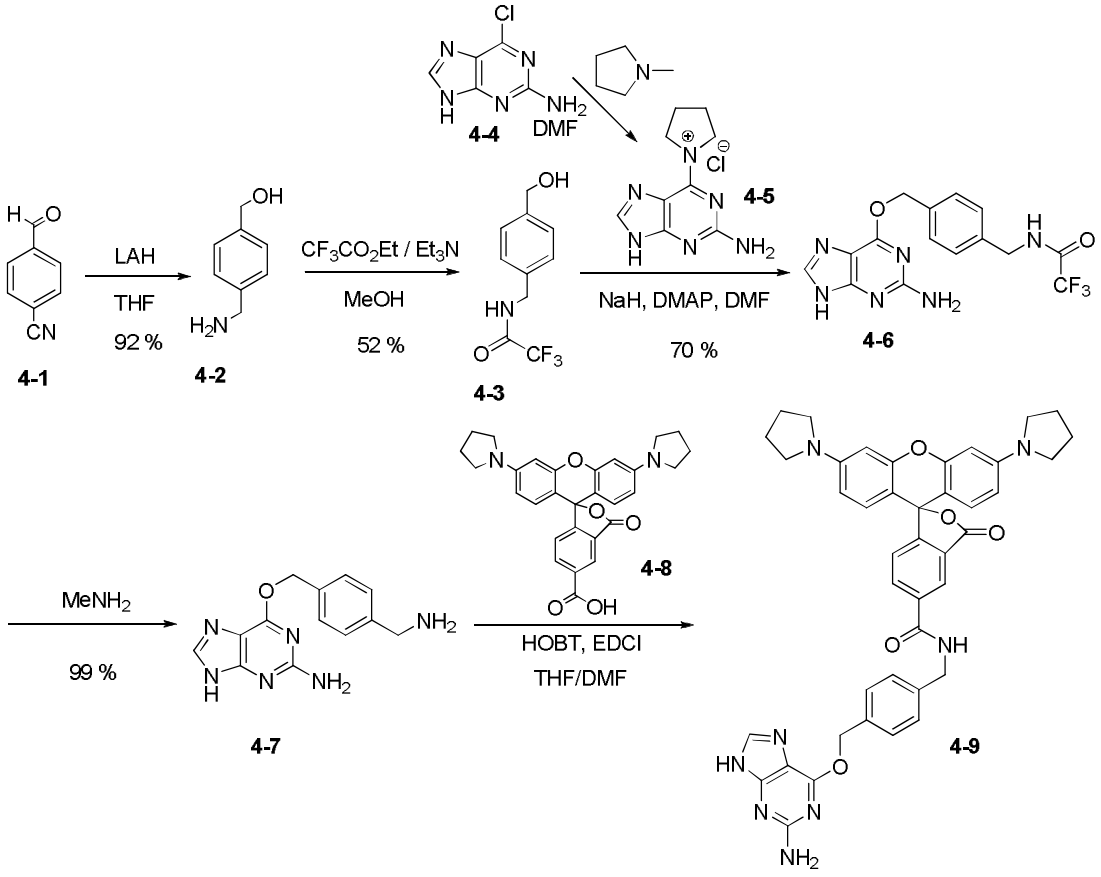


Figure 4-20. Thrombin cleavage of His-tag of MTS-AcGFP-AGT. All utilized concentrations led to complete cleavage.

4.3.1.2 Synthesis of Model Compound

Labeling of AGT based fusion proteins requires an alkyl guanine or benzyl guanine moiety. Therefore the model compound rhodamine was linked to a benzyl guanine group (Scheme 4-1 and Figure 4-18). The synthesis commenced with the reduction of 4-cyanobenzaldehyde (**4-1**) by $LiAlH_4$ to the corresponding alcohol and amine yielding **4-2**.[21] The amino group was protected as a trifluoroacetamide **(4-3)** and reacted with guanine derivative **4-5** to obtain **4-6**.[11] The resulting O^6-guanine derivative was deprotected in order to obtain **4-7**. Benzyl guanine (**4-7**) was coupled with rhodamine derivative **4-8**[22] affording the desired benzyl guanine rhodamine **4-9** (BGRH).



Scheme 4-1.

4.3.1.3 Labeling of AGT Based Proteins with Rhodamine Derivatives

Isolated protein MTS-AcGFP-AGT was incubated with rhodamine derivative **4-9** (BGRH) and the labeling reaction monitored by fluorescence spectroscopy (Figure 4-21). The comparison of the emission ([M]) and excitation ([X]) spectra of MTS-AcGFP-AGT before and after reaction with BGRH shows the formation of a FRET pair, thus indicating a successful labeling reaction (Figure 4-22 and Figure 4-23). Spectrum **D** shows the decrease of the fluorescence of the GFP domain and an energy transfer to the rhodamine label and subsequent emission of light. The best indicator for the generation of a FRET pair is the excitation spectrum **F** (Figure 4-24). A strong GFP component (see spectrum **B**) can now be observed at the far-red emission wavelength of 600 nm. Based on the observed changes of the fluorescence spectra the estimated FRET efficiency was about 44 % (for calculations see chapter 4.8.4). The Foerster radius (distance of 50 % FRET efficiency) was 60 Å. Therefore the distance between donor and acceptor was 64 Å. The distance of the fluorophore of GFP to its C-terminus is about 26 Å and the distance from the active site of AGT to its N-terminus is approximately 17 Å (based on the protein crystal structures). Adding 5 Å due to the size of the rhodamine label sums up to approximately 50 Å (the linker also has to be taken into account but its exact size is unknown). This estimation is in agreement with the determined donor - acceptor (DA) distance derived from the spectral data. Due to the strong distance dependence of the energy transfer efficiency the DA distance has to be in the narrow range of 0.5 R_0 and $2R_0$ (R_0 = Foerster radius) for meaningful measurements. Distances below 0.5 R_0 lead to almost complete donor quenching while distances of more than $2R_0$ result in inefficient energy transfer. The fusion protein MTS-AcGFP-AGT, labeled with rhodamine, has a DA

distance close to R_0 and hence represents an excellent starting point for attaching colored metal chelators.

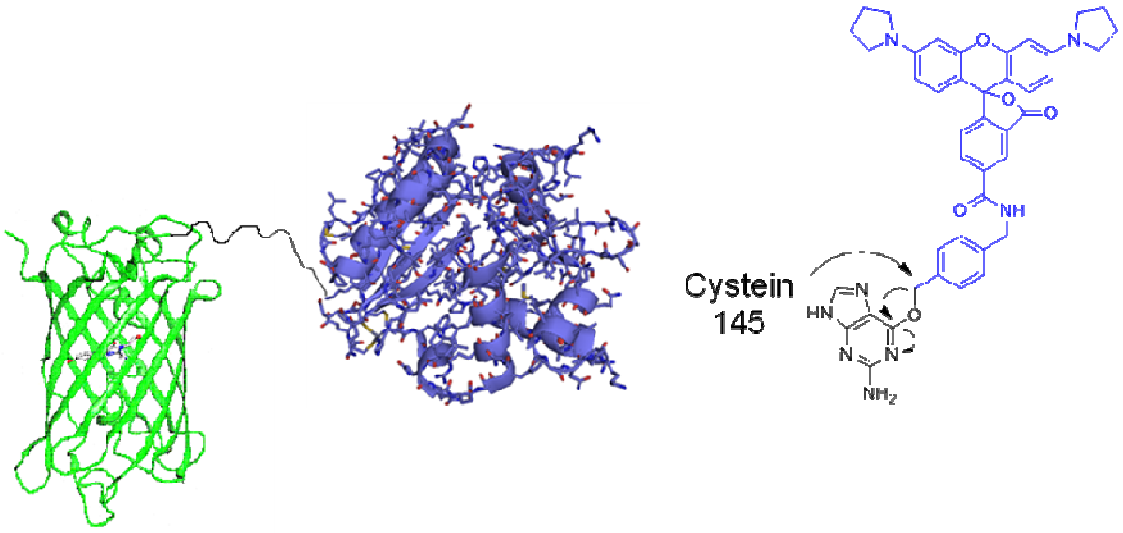


Figure 4-21. Labeling of MTS-AcGFP-AGT with rhodamine.

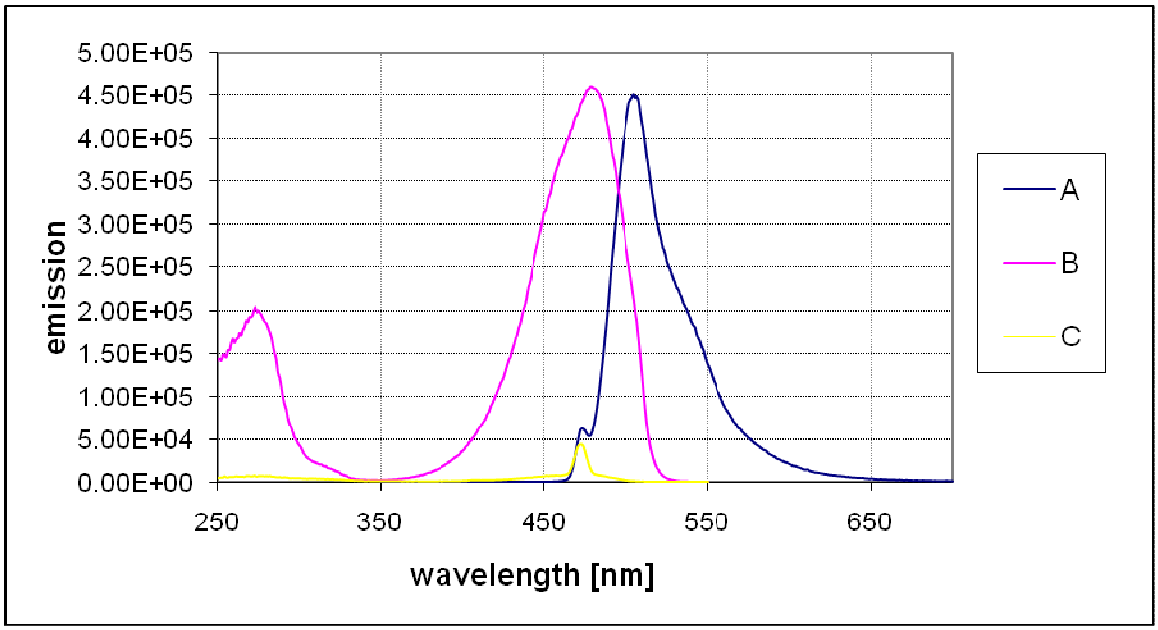


Figure 4-22. Fluorescence spectrum of unreacted MTS-AcGFP-AGT (0.29 µM in 50 mM phosphate buffer, pH 8.0). **A.** [M] MTS-AcGFP-AGT (exc.: 475 nm). **B.** [X] MTS-AcGFP-AGT (emm.: 505 nm). **C.** [X] MTS-AcGFP-AGT (emm.: 470 nm).

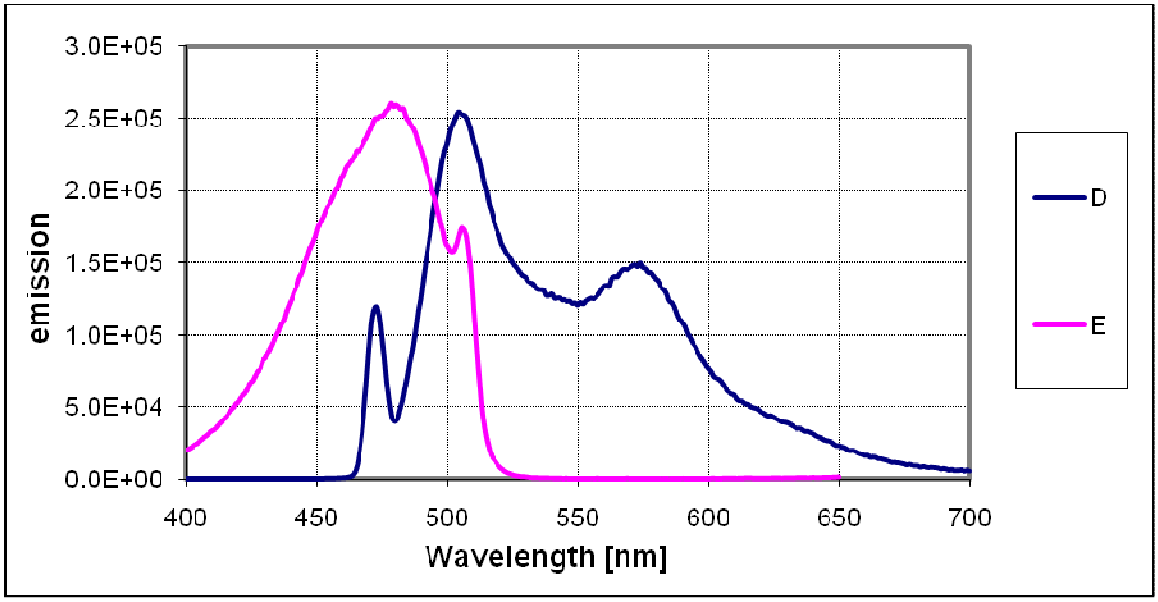


Figure 4-23. Fluorescence spectrum of the labeling reaction of MTS-AcGFP-AGT with BGRH (0.29 µM each in 50 mM phosphate buffer, pH 8.0). **D.** [M] MTS-AcGFP-AGT + BGRH (exc.: 475 nm). **E.** [X] MTS-AcGFP-AGT + BGRH (emm.: 505 nm).

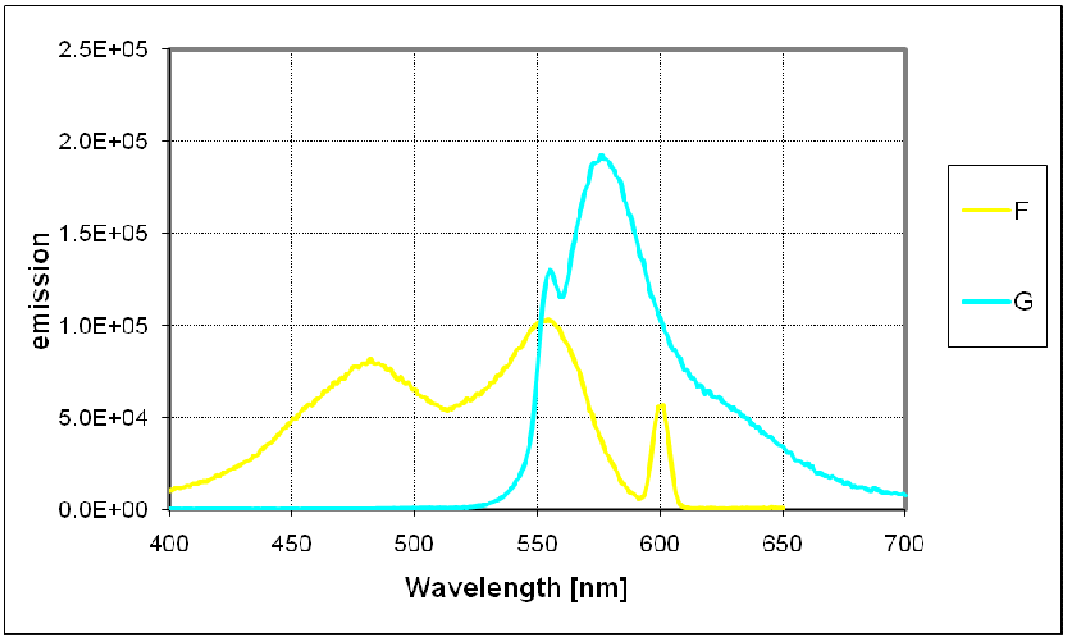


Figure 4-24. Fluorescence spectrum of the labeling reaction of MTS-AcGFP-AGT with BGRH (0.29 µM each in 50 mM phosphate buffer, pH 8.0). **F.** [X] MTS-AcGFP-AGT + BGRH (emm.: 600 nm). **G.** [M] MTS-AcGFP-AGT + BGRH (exc.: 554 nm).

The formation of a FRET pair was consistent with the observed change of fluorescent lifetimes (Figure 4-25). The change of the fluorescence intensities of the donor in case of energy transfer is directly proportional to the change of fluorescence lifetimes. The FRET efficiency was 44 % therefore the fluorescence lifetime should also decrease by 44 %. Indeed, the fluorescence lifetime dropped from approximately 2.7 ns to 1.5 ns, which is a decrease of 45 % and close to the measured decrease of 44 % of the fluorescence intensity of the donor (fluorescence lifetimes analyzed by timemaster).

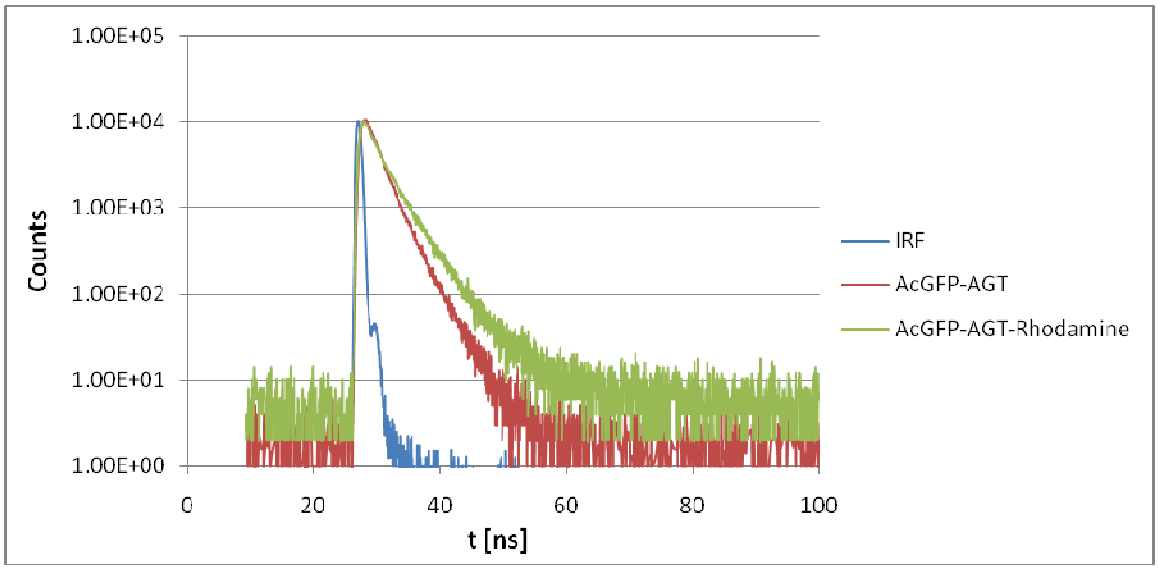


Figure 4-25. Fluorescence decay of labeled and unlabeled MTS-AcGFP-AGT.

4.3.2 Labeling of Split-Intein Based Fusion Proteins

4.3.2.1 Expression and Isolation of Split-Intein Based Fusion Proteins

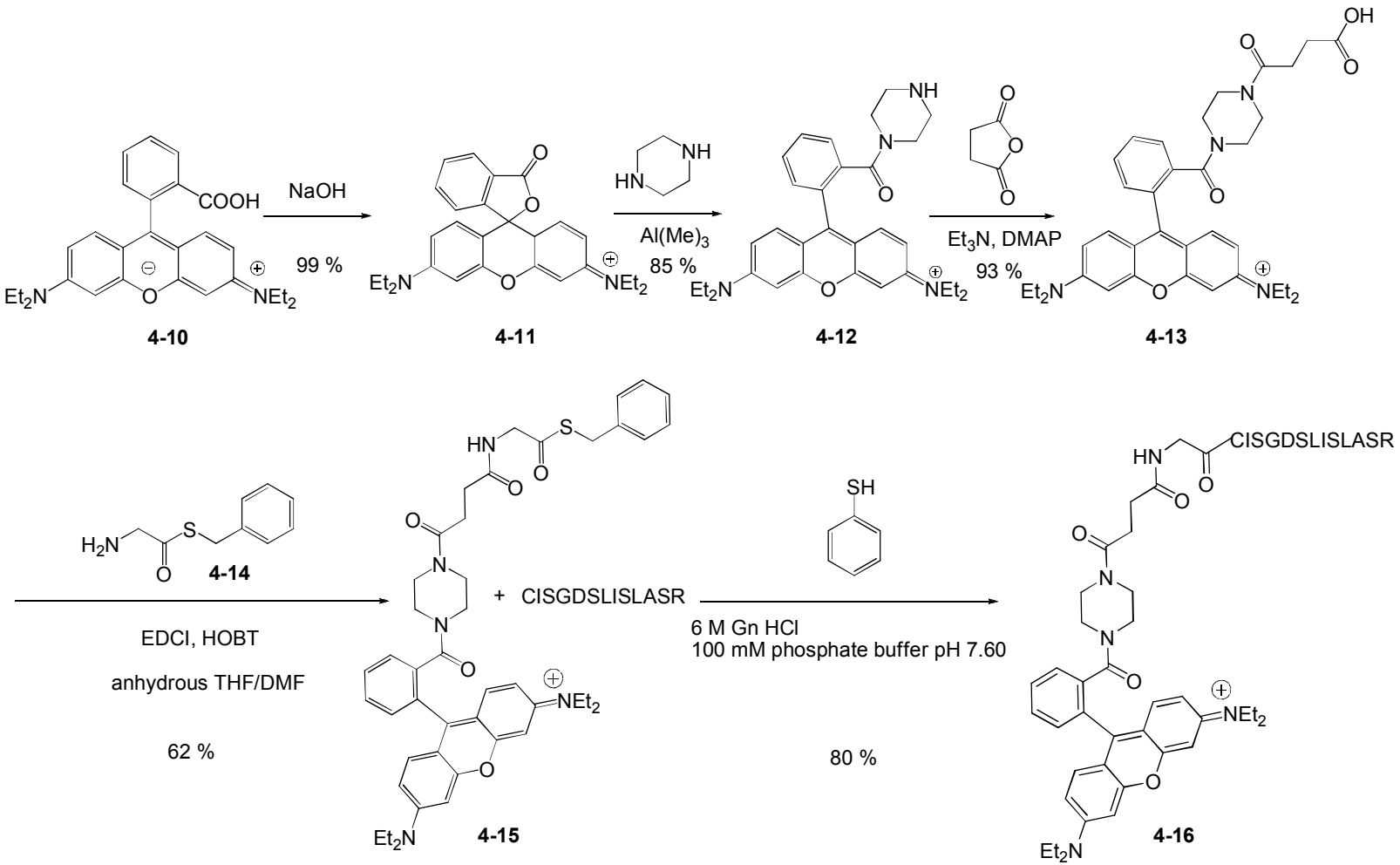
E. coli bacteria were transformed with plasmids for bacterial expression encoding intein^{C}-AcGFP-TGN38 and intein^{C}-Cerulean-TGN38 (plasmids **Q** and **R**). No formation of the fusion proteins was observed after induction of expression followed by lysis of the bacterial cells (data not shown). The expression of TGN38 in bacteria has not been reported. Eukaryotic membrane and glycosylated proteins usually are difficult to express in bacteria because bacteria do not glycosylate proteins. Glycosylation is important for protein folding and stability in mammalian cells. Hence, the failure to express these heavily glycosylated membrane proteins in *E. coli* was not surprising. It was attempted to circumvent this problem by exporting the fusion protein to the periplasm by inserting an export signal sequence to periplasm (plasmids not shown). This strategy often leads to

the successful expression of eukaryotic because the periplasmic medium is much more similar to the eukaryotic intracellular medium than the bacterial cytoplasmic medium; however, this strategy did not lead to the desired fusion proteins either (data not shown). Since TGN38 is only required to direct the fusion protein to the TGN the design was simplified to intein^{C}-AcGFP and intein^{C}-Cerulean (plasmids **S** and **T**), which is sufficient to examine trans-splicing reactions *in vitro*. The encoded His-tagged protein was successfully expressed and purified (Mw ≈ 47 kDa) (Figure 4-27 control); however, the purified protein solution contained a significant amount of a protein with a molecular weight of about 32 kDa. This was probably His-tagged AcGFP, generated by ribosomes initiating translation at the remaining start codon of AcGFP instead of at intein^{C}, and was not expected to interfere with the *trans*-splicing reaction.

4.3.2.2 Design and Synthesis of Split-Intein^{N} Containing Model Compound

Labeling of split-intein based fusion proteins requires a split-intein moiety. Hence, the model compound rhodamine was linked to an 11 amino acid split-intein^{N} moiety (Scheme 4-2 and Figure 4-26). The polypeptide also contains the necessary N-extein. The synthesis commenced with the formation of the lactone **4-11** by the reaction of rhodamine B (**4-10**) with NaOH.[23] Lactone **4-11** was transformed to the tertiary amide **4-12** by piperazine and $Al(Me)_3$. The secondary amine of the piperazyl moiety (**4-11**) was reacted with succinic anhydride, yielding carboxylic acid **4-13**. Amide coupling with the synthesized **4-14** (synthesis not shown in Scheme 4-2), catalyzed by EDCI and HOBT, led to the rhodamine thioester **4-15**. The final rhodamine derivative **4-16** was obtained by

native chemical ligation with the polypeptide, containing an N-terminal cysteine, and a large excess of thiophenol.


NaOH
99 %
Al(Me)3
85 %
Et3N, DMAP
93 %
4-10
4-11
4-12
4-13
4-14
EDCI, HOBT
anhydrous THF/DMF
62 %
+ CISGDSLISLASR
4-15
6 M Gn HCl
100 mM phosphate buffer pH 7.60
80 %
CISGDSLISLASR
4-16

Scheme 4-2.

4.3.2.3 Labeling of Split-Intein Based Fusion Proteins with Rhodamine Derivatives

It was attempted to label the protein intein^{C}-AcGFP by *trans* splicing with split-inteins utilizing rhodamine derivative **4-16** (Figure 4-26). This method was based on the labeling of the protein thioredoxin with fluorescein by Mootz et al.[8] Successful labeling should generate a FRET pair similar to MTS-AcGFP-AGT-rhodamine (chapter 4.3.1); however no change was observed by fluorescence spectroscopy (data not shown). This was confirmed by SDS-PAGE (Figure 4-27). *Trans*-splicing should result in the formation of spliced out intein^{C} (Mw ≈ 17 kDa). The missing band at 17 kDa clearly showed that no splicing reaction occurred.

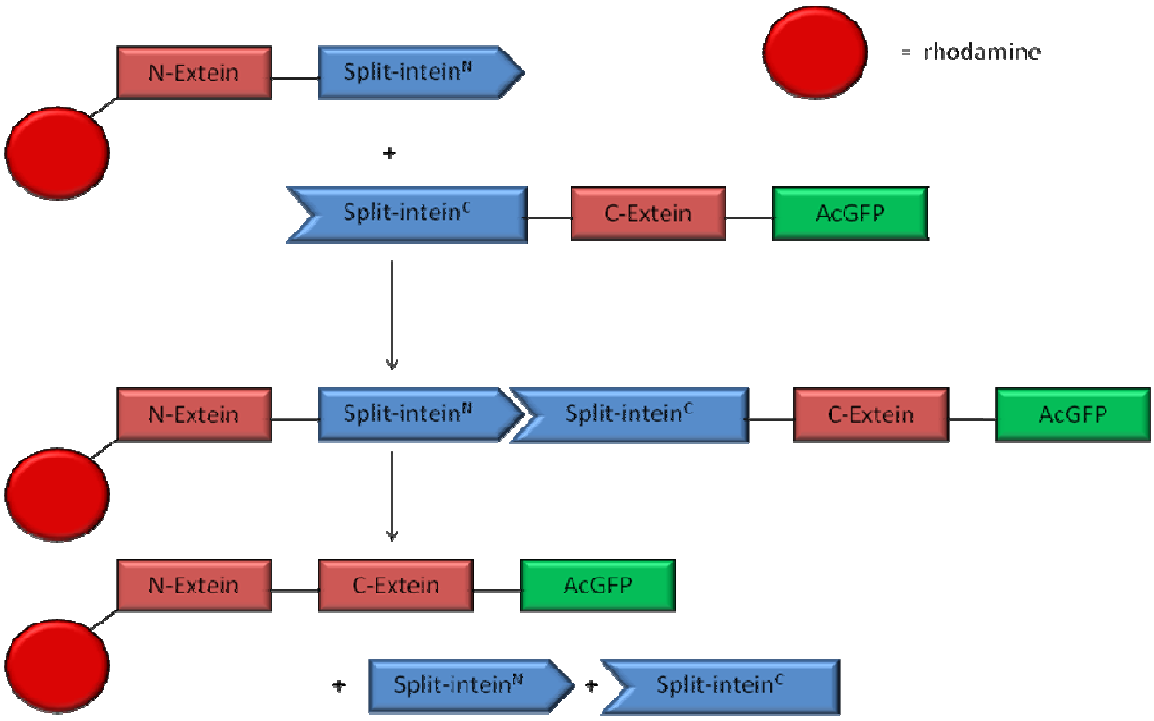


Figure 4-26. Labeling of AcGFP with rhodamine by *trans* splicing with split-inteins. The splicing reaction is induced by recombination of the split-inteins and yields labeled protein and the excised split-inteins.

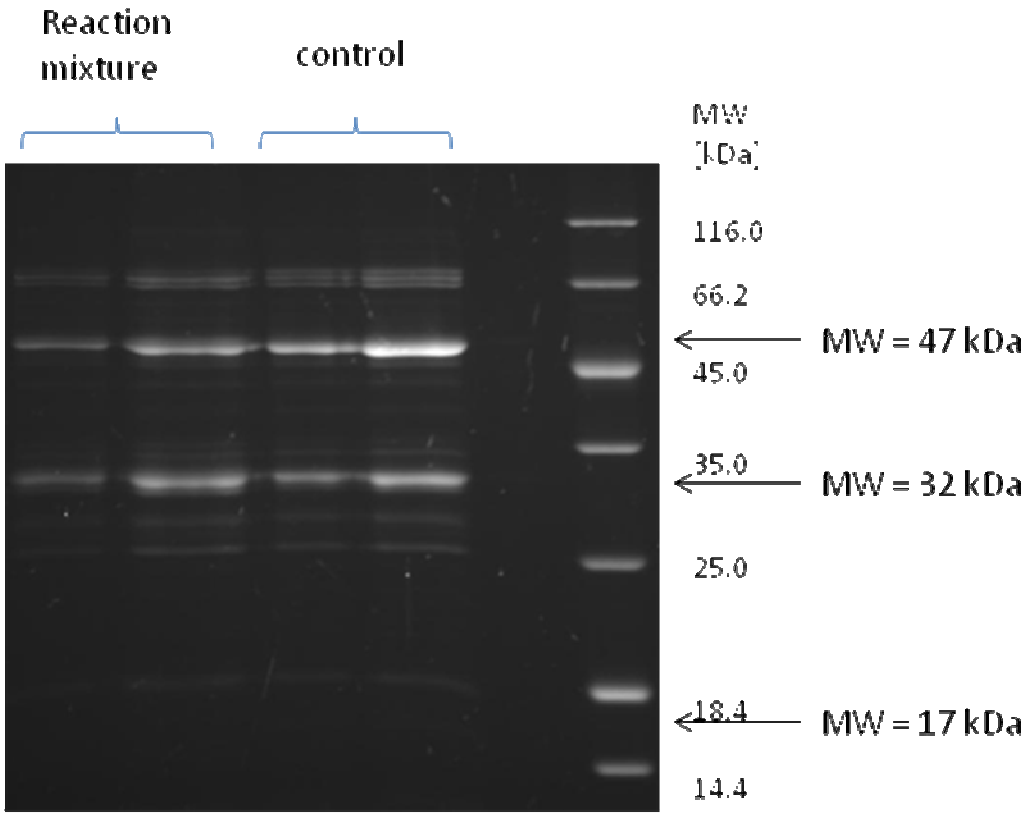


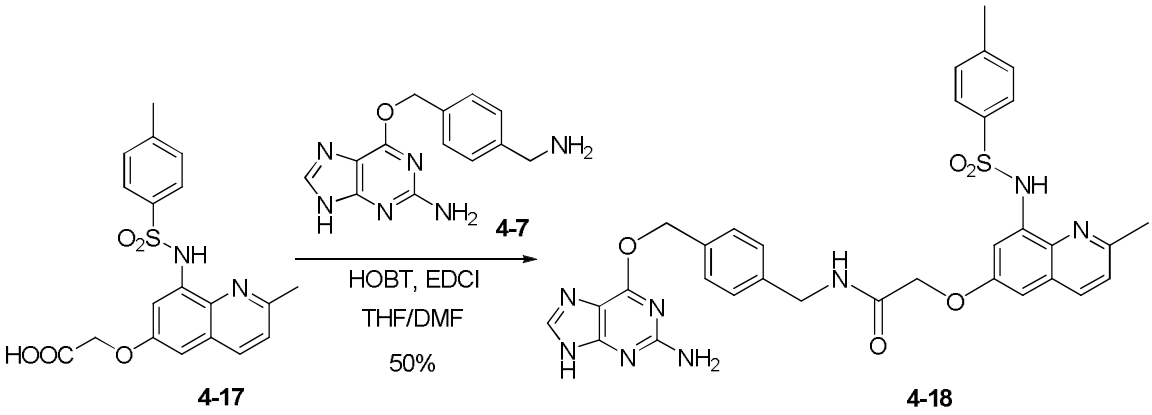
Figure 4-27. SDS-PAGE gel of the trans-splicing reaction with split inteins. Intein^{C}-AcGFP can be seen at 47 kDa and an impurity at 32 KDa that is most likely His-tagged AcGFP. No band of excised intein^{C} was observed at 17 kDa suggesting that *trans*-splicing did not occur.

4.4 Labeling of Proteins with Zinc Chelators

In order to generate semi-synthetic proteins for the sensing of Zn-ions suitable colored or fluorescent Zn-chelators needed to be synthesized (for requirements of the metal chelators see chapter 4.1.3).

4.4.1 Synthesis of Benzyl Guanine Zinquin

Benzyl guanine zinquin **4-18** was synthesized by HOBT and EDCI catalyzed coupling of zinquin (**4-17**) with benzyl guaine (**4-7**) (Scheme 4-3).



Scheme 4-3.

4.4.2 Labeling of Proteins with Benzyl Guanine Zinquin

Benzyl guanine zinquin showed excellent solubility in CH_2Cl_2 (which is somewhat surprising because zinquin and benzyl guanine are poorly soluble in CH_2Cl_2); however, **4-18** exhibited poor solubility in water and only fair solubility in DMF and DMSO. The poor solubility in water probably derived from aggregation due to hydrophobic effects (see chapter 1.1.2) and prevented the examination of properties such as binding affinities and binding kinetics of the zinc chelator. The labeling of AGT based fusion proteins was sluggish because precipitation occurred quickly after addition of benzyl guanine zinquin to the reaction mixture despite adding 10 % of DMF (v/v) to the reaction mixture. Hence, **4-18** was not suitable as a zinc chelator for labeling proteins with the goal of *in vivo* cell imaging.

4.4.3 Labeling of Proteins with Alternative Zinc Chelators

Due to the unsatisfying properties of benzyl guanine zinquin **4-18** alternative benzyl guanine zinc chelators were developed. The Fahrni research group developed a fluorescent, selective and ratiometric Zn(II) chelator that is suitable for metal sensing in methanol (Figure 4-28).[24] This dye, designed for two-photon absorption (also suitable for one photon absorption), has attractive properties such as a binding affinity in the µM range but is not suitable for metal sensing in aqueous medium. Ratiometric behavior was not observed in aqueous medium, probably due to processes such as aggregation, dimerization or excimer formation (data not shown). Attachment to a protein would solve these problems because these processes cannot occur if the molecules are separated from each other. That is why a benzyl guanine derivative of this type of dye was synthesized (Figure 4-29).[25]

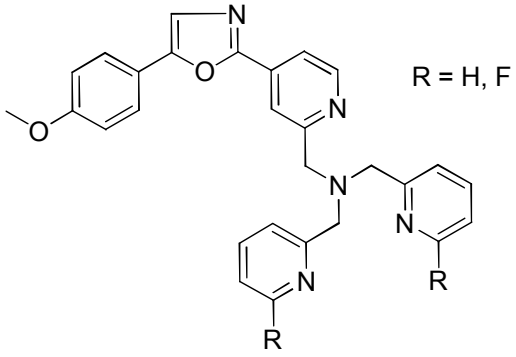


Figure 4-28. Emission ratiometric Zn(II) sensor. The sensor contains a tris(picolyl) metal-chelating unit with an anisole donor connected by a rigid oxazole core.[24]

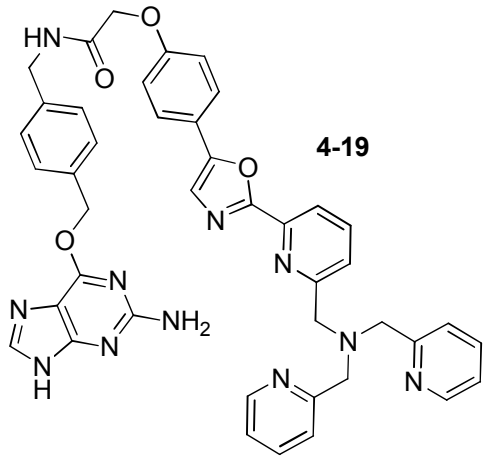


Figure 4-29. Benzyl guanine modified Zn(II) emission ratiometric sensor (see Figure 4-25). [24]

The fluorescence properties of **4-19** in methanol and in aqueous buffered solution differ significantly (Figure 4-30). In methanol **4-19** shows basically the same behavior as the Zn(II) chelator without the benzyl guanine moiety. Upon excitation (at 350 nm) of **4-19** in methanol in the absence of Zn(II) (TPEN was added to remove trace amounts of Zn(II)) fluorescence emission was observed with an intensity maximum at 418 nm (**A**). Saturation of the zinc chelator **4-19** with Zn(II) resulted in a large spectral shift and fluorescence enhancement (**B**). In contrast the fluorescence intensity of **4-19** is much weaker in aqueous buffer (in the absence of Zn(II)) and the shape of the curve is not Gaussian (bell-shaped) suggesting the presence of more than one emissive state (**C**). Saturation of the zinc chelator **4-19** with Zn(II) in aqueous buffer exhibited a fluorescence spectrum (**D**) similar to Zn(II) saturated **4-19** in methanol.

The expressed and purified protein MTS-AGT was labeled with **4-19** and the reaction monitored by fluorescence spectroscopy (Figure 4-31). Reaction of **4-19** with MTS-AGT led to a considerable change of the fluorescence emission (compared to **C**)

already after five minutes (**E**). The intensity of the fluorescence emission kept increasing and reached its maximum after about one hour (**F**). The resulting spectrum illustrated a remarkable similarity to the “zinc-free” fluorescence spectrum of **4-19** in methanol (**A**). This suggests that the protein was labeled with the zinc(II)-chelator, thus inhibiting processes such as aggregation, dimerization or excimer formation. Upon saturation of the zinc-chelator domain of the formed semi-synthetic protein with Zn^{2+} a fluorescence emission was exhibited (**G**) similar to Zn(II)-saturated **4-19** in aqueous buffer (**D**); however, the emission maximum (474 nm (**G**) and 484 nm (**D**)) was shifted by 10 nm. This shift is another indication for the successful labeling of the protein MTS-AGT. The semi-synthetic protein formed by the *in vitro* reaction of MTS-AGT and **4-19** represents, due to its fluorescence properties in the presence and absence of Zn^{2+}, a very useful platform for the goal of *in vivo* cell imaging of zinc-ions. Therefore the focus shifted on the development of chelators for iron and copper, suitable for protein labeling.

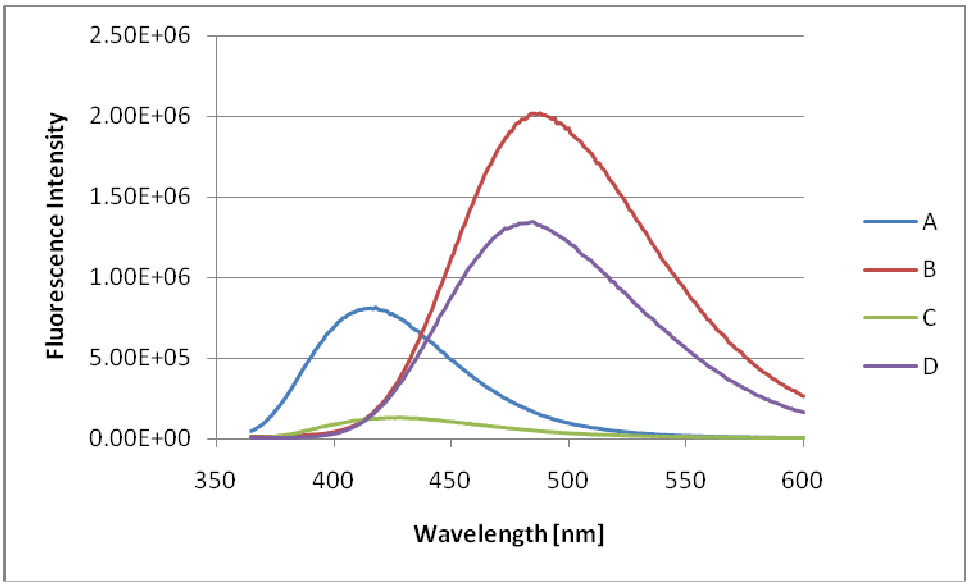


Figure 4-30. Fluorescence emission spectra (excitation at 350 nm) of **4-19** (4 µM) in methanol or aqueous buffered solution (10 mM PIPES, 10 mM HEDTA, pH 7.06, 0.1 M KCl). **A. 4-19** + TPEN in MeOH. **B. 4-19** + Zn^{2+} in MeOH. **C. 4-19** in aqueous buffer. **D. 4-19** + Zn^{2+} in aqueous buffer.

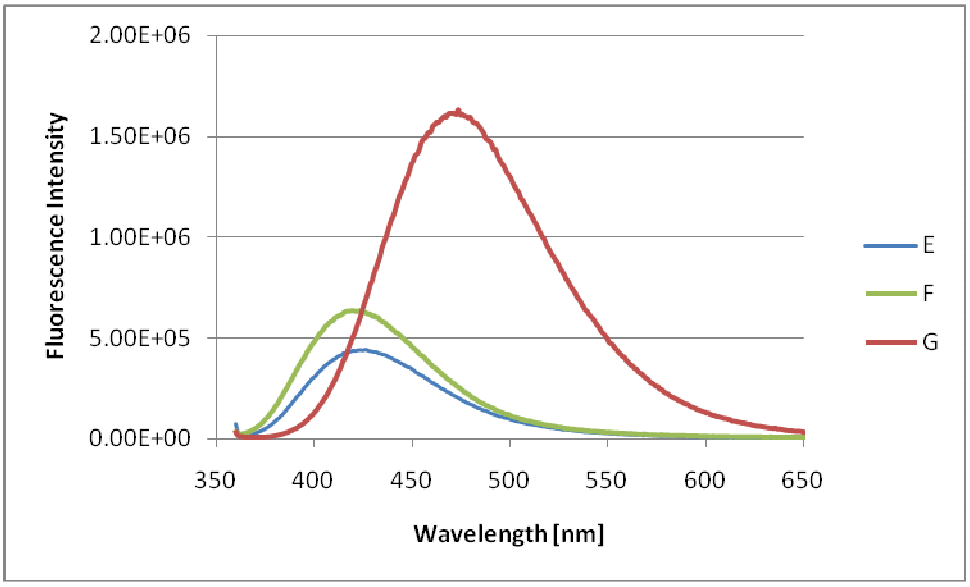
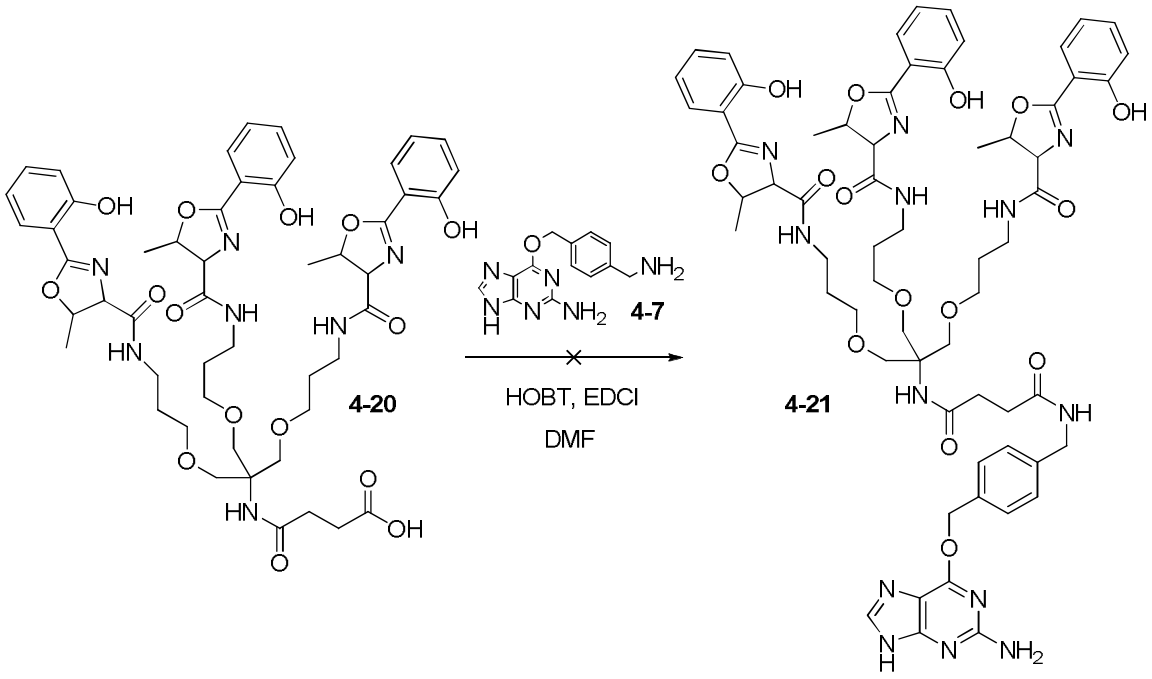


Figure 4-31. Fluorescence emission spectra (excitation at 350 nm) of **4-19** (6 µM) reacted with MTS-AGT (6 µM) in aqueous buffered solution (10 mM PIPES, 10 mM HEDTA, pH 7.06, 0.1 M KCl, 1 mM DTT). **E.** reaction time = 5 min. **F.** reaction time = 1 h. **G.** F + Zn^{2+}.

4.5 Labeling of Proteins with Iron Chelators

In order to generate semi-synthetic proteins for the sensing of Fe-ions suitable colored or fluorescent Fe-chelators needed to be synthesized. Kikkeri et al. have recently reported on the synthesis of a tripodal phenol-oxazoline based Fe(III) chelator that exhibits a red color upon metal binding (**4-20** in Scheme 4-4).[26] Therefore it was attempted to synthesize a benzyl guanine derivative in the same fashion as for the Zn(II)-chelators (see chapter 4.4). Unfortunately, all efforts to synthesize benzyl guanine derivative **4-21** failed. Consequently, AGT based fusion proteins could not be labeled with any iron chelators.



Scheme 4-4.[27]

4.6 Labeling of Proteins with Copper Chelators

In order to generate semi-synthetic proteins for the sensing of Cu(I)-ions suitable colored or fluorescent Cu(I)-chelators needed to be synthesized. The copper chelator has to form 1:1 complexes; however, no colored or fluorescent chelators have been reported that bind Cu(I) in a 1:1 fashion. Therefore it was attempted to synthesize 1:1 binding Cu(I)-chelators based on the known copper chelators bicinchoninic acid and bathocuproine, which bind Cu(I) in a 2:1 fashion (Figure 4-32 and 4-33). Based on molecular modeling a diethanolamine or resorcinol (1,3-benzenediol) based linker would allow the ligands to rotate to tetrahedral metal binding position, thus forming a complex with 1:1 Cu(I)-ligand stochiometry.

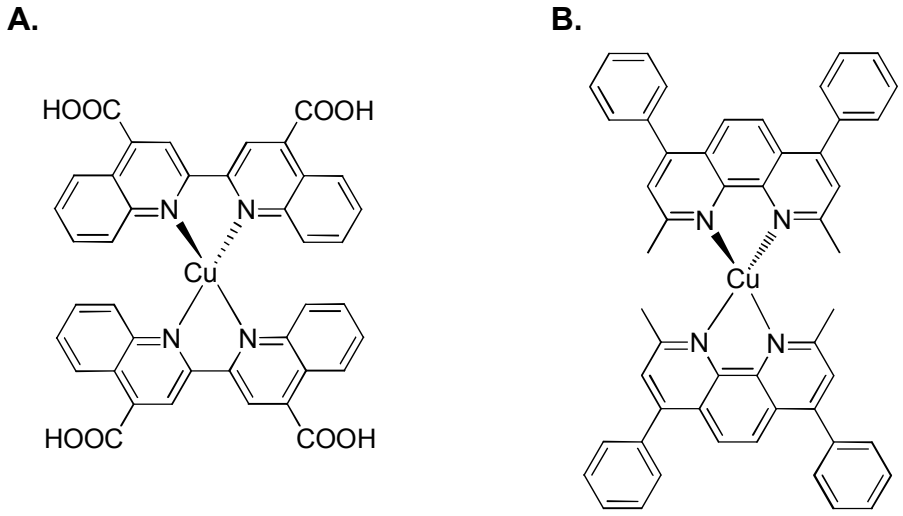


Figure 4-32. Tetrahedral 2:1 Cu(I) complexes of bicinchoninic acid (**A**) and bathocuproine (**B**). Both metal complexes are colored (**A** = purple; **B** = red).

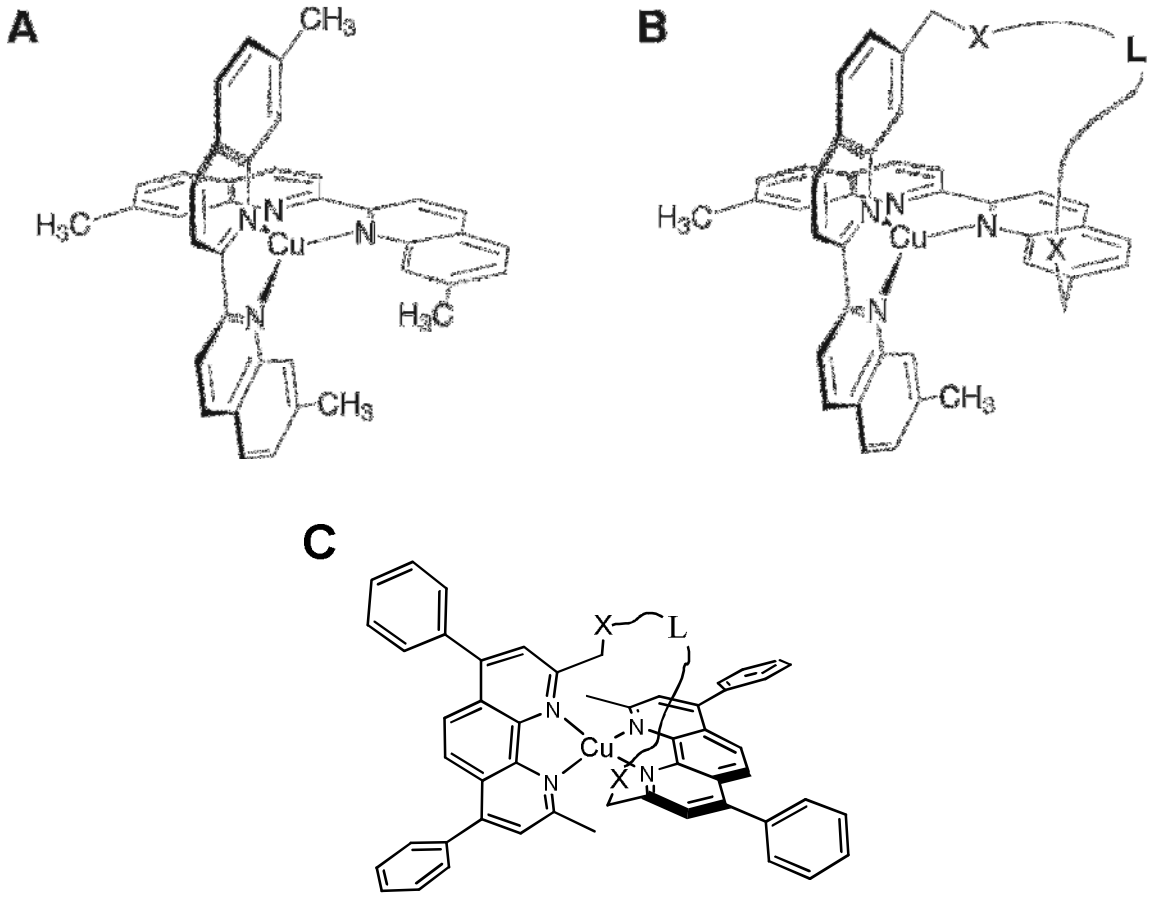


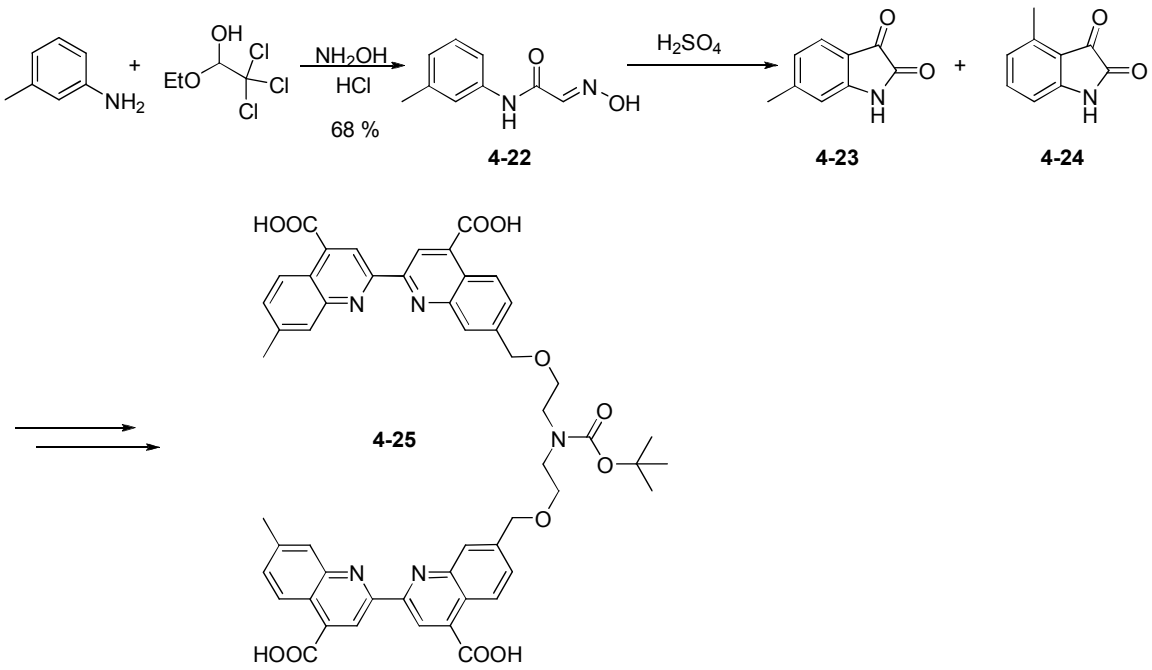
Figure 4-33. Structures of Cu(I) complexes of bicinchoninic acid and bathocuproine (generated by molecular modeling with MM2 force field calculations).[28] **A.** Dimethylbicinchoninic acid complex with 2:1 ligand-metal stochiometry. **B.** Dimethylbicinchoninic acid complex connected by a linker with 1:1 ligand-metal stochiometry. **C.** Bathocuproine complex connected by a linker with 1:1 ligand-metal stochiometry.

4.6.1 Synthesis of Bicinchoninic Acid Based Derivatives

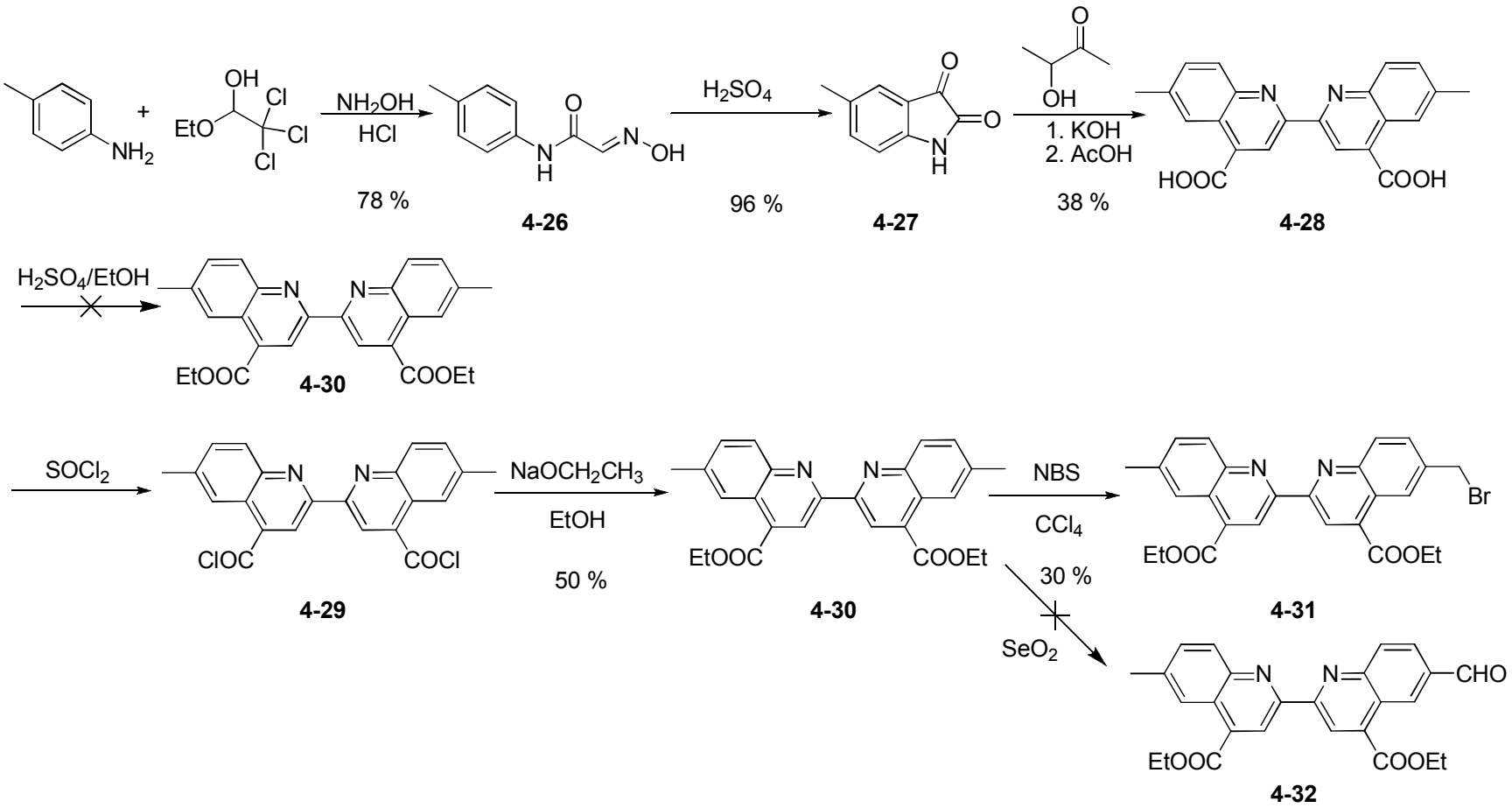
The possibly 1:1 binding Cu(I) chelator **4-25** was attempted to synthesize starting with the formation of *iso*nitrosoacetanilide **4-22** by the reaction of *m*-toluidine and 2,2,2-trichloro-1-ethoxy ethanol (Scheme 4-5).[29] Reaction to the corresponding isatin derivative by concentrated sulfuric acid resulted in the formation of a mixture of 4-methyl

isatin **4-24** and the desired 6-methyl isatin **4-23**.[30] The purification of **4-23** requires several laborious crystallization and precipitation steps,[31-33] therefore the target molecule and consequently the reactants and intermediates were altered (Scheme 4-6). Utilizing *p*-toluidine as starting material resulted avoided the formation of mixtures of isatin isomers; thus, 5-methyl isatin **4-27** was obtained in good yield (after synthesis of *iso*nitrosoacetanilide **4-26**). The reaction of **4-26** with KOH and acetoin (Pfitzinger reaction), followed by acidification, yielded the dimethyl bicinchoninic acid derivative **4-28**.[34] Acid catalyzed acidification did not result in the formation of the ethyl ester **4-30**. That is why 4-28 was transformed to an acyl chloride (**4-29**) and reacted with sodium ethoxide in order to obtain the ethyl ester **4-30**. Radical bromination by a stochiometric amount of NBS provided the monobromo derivative **4-31**. Attempts to couple two molecules of bicinchoninic acid ester **4-31** with the *t*-boc protected diethanolamine through the formation of alkyl oxides by NaH failed (Scheme 4-7).

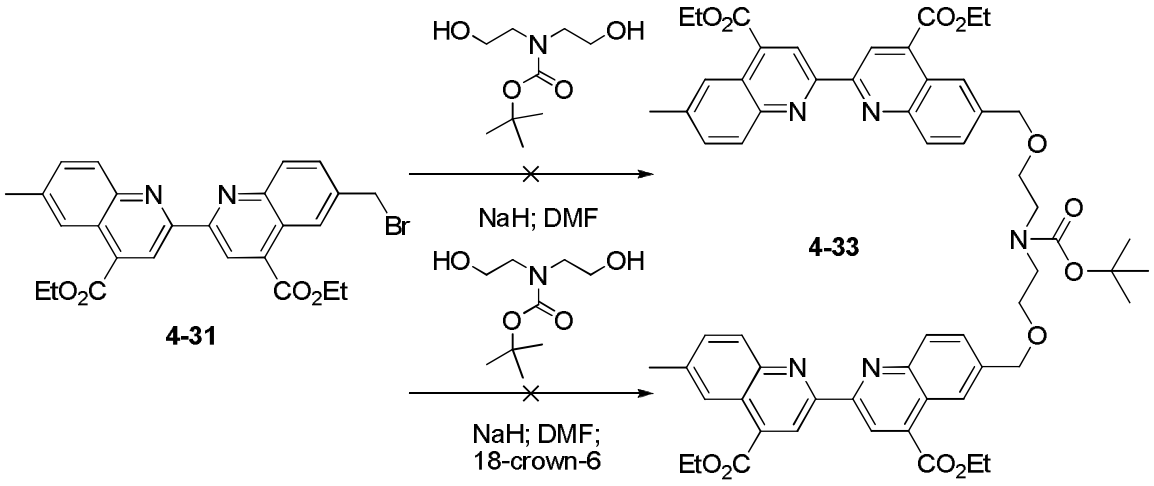
Since small amounts of water can lead to the formation of NaOH, which can hydrolyze the ester moiety, the more base-stable *t*-butyl ester derivatives of bicinchoninic acid were synthesized (Scheme 4-8). Esterification of acyl chloride **4-29** with K-*t*BuO did not succeed. Hence, dimethyl bicinchoninic acid **4-28** was transformed to the corresponding *tert*-butyl ester (**4-34**) with DMAP and EDCI. Radical bromination with NBS provided the monobromo derivative **4-35**; however, coupling of two bicinchoninic acid molecules (**4-35**) in order to yield **4-36** failed.


NH2
OH
EtO
Cl
Cl
Cl
NH2OH
HCl
68 %
4-22
H2SO4
4-23
4-24
HOOC
COOH
N
N
O
N
O
O
4-25
N
N
HOOC
COOH

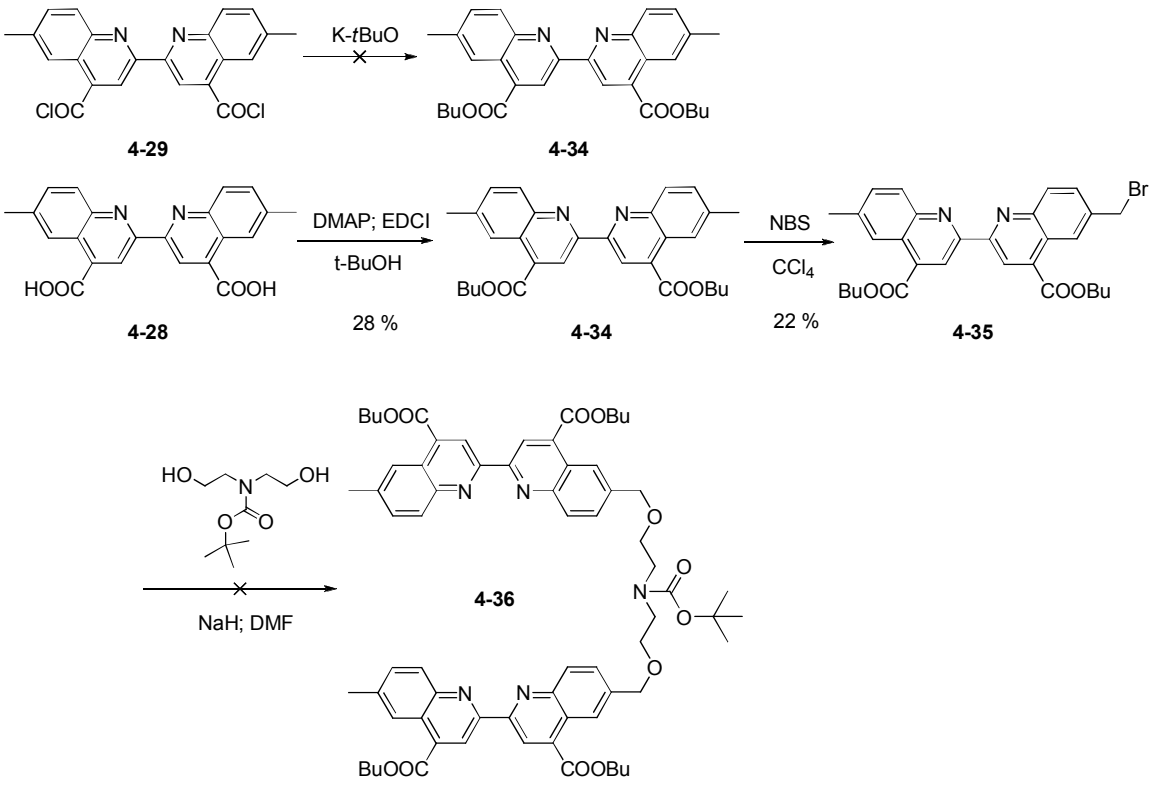
Scheme 4-5.


NH2
OH
Cl
EtO
Cl
Cl
NH2OH
HCl
78 %
4-26
H2SO4
96 %
4-27
O
OH
1. KOH
2. AcOH
38 %
HOOC
COOH
4-28
H2SO4/EtOH
EtOOC
COOEt
4-30
SOCl2
ClOC
COCl
4-29
NaOCH2CH3
EtOH
50 %
EtOOC
COOEt
4-30
NBS
CCl4
30 %
Br
4-31
SeO2
CHO
4-32

Scheme 4-6.



Scheme 4-7.

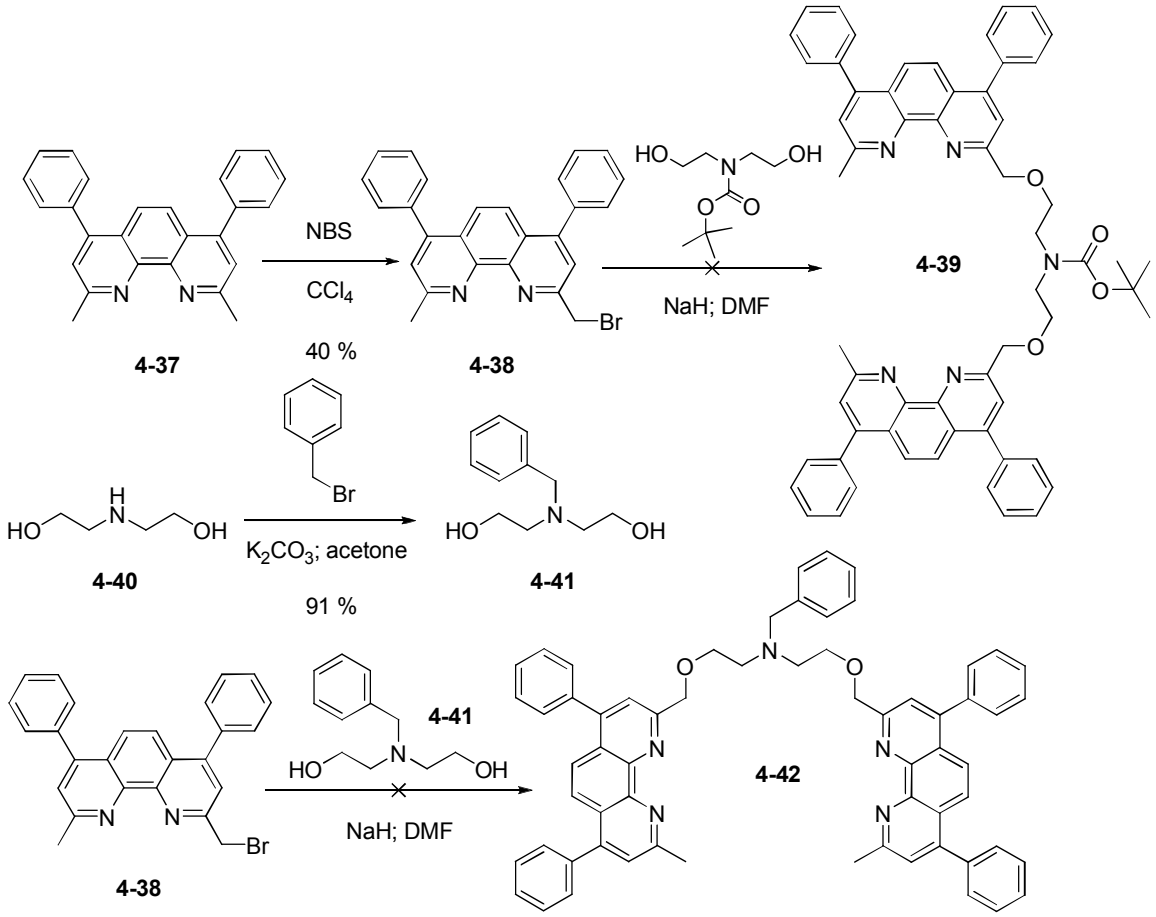


Scheme 4-8.

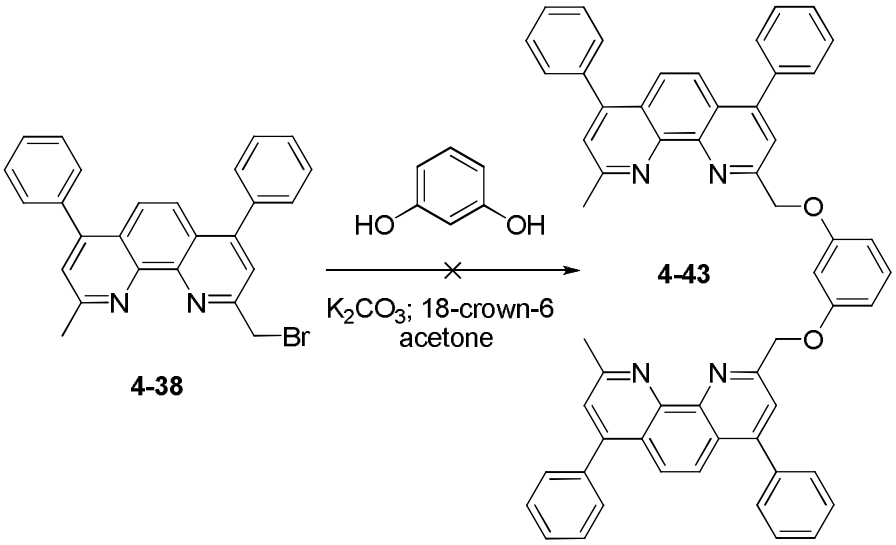
4.6.2 Synthesis of Bathocuproine Based Derivatives

Due to the failure to form a 1:1 binding Cu(I) chelator based on bicinchoninic acid the focus shifted to utilizing bathocuproine derivatives (Scheme 4-9). Bromination of bathocuproine (**4-37**) was achieved by radical bromination with NBS yielding **4-38**. Similar to the bicinchoninic acid, the coupling of two molecules of bathocuproine in order to obtain **4-39** failed. Therefore, a different linker was utilized (**4-41**). The benzyl protected linker **4-41** did not result in the formation of the desired connected bathocuproine derivative **4-42**.

Aromatic alcohols usually require milder reaction conditions than alkyl alcohols for reactions with halides. That is why monobromo bathocuproine **4-38** was reacted with resorcinol (Scheme 4-10). The yield of the reaction was insufficient (< 1 % based on mass spectrometry); however, the reaction proceeded in a “clean” fashion, producing mostly unreacted starting material, bathocuproine linked to one alcohol of resorcinol, and the desired product **4-43** (determined by mass spectrometry and NMR spectroscopy). Therefore this approach is a good starting point for future attempts to synthesize bicinchoninic acid or bathocuproine based 1:1 binding Cu(I)-chelators.


NBS
CCl4
40 %
4-37
4-38
HO
OH
N
O
O
NaH; DMF
4-39
Br
4-40
K2CO3; acetone
91 %
4-41
4-42

Scheme 4-9.



Scheme 4-10.

4.7 Conclusion

Plasmids encoding fluorescent proteins, targeting sequences and AGT or intein fusion domains for eukaryotic and prokaryotic expression were generated. Successful targeting of intracellular compartments (mitochondria, nucleus and TGN) was confirmed by light microscopy experiments with transfected mammalian cells. The encoded proteins were overexpressed in *E.coli* and purified. *In vitro* labeling experiments of the fusion proteins with rhodamine derivatives succeeded with AGT based fusion proteins resulting in the formation of a FRET-pair; however, labeling of fusion proteins by *trans*-splicing with split-inteins failed. Zinc(II)-chelator **4-19** was attached to an AGT based protein and the resulting semi-synthetic protein exhibited strong changes of fluorescence in the presence of zinc(II). This represents an important step towards the goal of in vivo cell imaging of labile zinc(II) pools. Iron chelators suitable for protein-labeling could not

be synthesized. Despite extensive efforts, all attempts failed to generate a 1:1 binding Cu(I)-chelator.

4.8 Experimental Section

4.8.1 Materials and Reagents and General Techniques

All starting materials, reagents and solvents were purchased from Aldrich (Milwaukee, WI), Fisher Scientific (Pittsburgh, PA) or VWR (West Chester, PA). NMR: δ in ppm versus SiMe4 (0 ppm, ^{1}H, 400 MHz). MS: selected peaks; m/z. Flash chromatography: Merck silica gel (240-400 mesh). TLC: 0.25 mm, Merck silica gel 60 F_{254}, visualizing at 254 nm or with permanganate solution (2 % $KMnO_4$, 5 % K_2CO_3). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). Ampicillin was used at 50 µg/mL and Kanamycin at 30 µg/mL concentration. Unless otherwise stated, standard protocols were used. All plasmids were verified by DNA sequencing. Mutated sites are in indicated in bold and red.

4.8.2 Generation of Plasmids

CMV-MTS-AcGFP1-AGT (A). The host or template vectors were the commercially available plasmids pSEMXT-26m (Covalys) containing AGT and pAcGFP1-Mito (Clontech) containing AcGFP1 and a mitochondrial targeting sequence. PAcGFP1-mito was digested by *Nhe*I and *Not*I and the isolated gene encoding MTS-AcGFP1 was cloned into the *Nhe*I and *Not*I sites of pSEMXT-26m. The resulting plasmid (pMito-AcGFP1-SNAP) contained an opal stop codon (TGA). Therefore pMito-AcGFP1-SNAP was used as a template for site-directed mutagenesis according to the QuickChange protocol (Stratagene) with the oligonucleotides 5'-GGCATGGATGAGCTGTACAAG**GCTCTA**GCCGCGATATCAAGCTTAC-3' and 5'-GTAAGCTTGATATCGCGGC**TAGAGC**CTTGTACAGCTCATCCATGCC-3' to give intermediate plasmid **pre-A-1**.

A second *Bsrg*I restriction site, within the AcGFP1 gene of **pre-A-1**, was removed by the introduction of a silent mutation by site-directed mutagenesis according to the QuickQhange protocol (Stratagene) with the oligonucleotides 5'-CTACAACGCCCACAATGTGTA**T**ATCATGACCGACAAGGCC-3' and 5'-GGCCTTGTCGGTCATGAT**A**TACACATTGTGGGCGTTGTAG-3' to yield intermediate plasmid **pre-A-2**. The obtained unique *Bsrg*I site allows straightforward replacement of the AcGFP gene by common GFP-derivatives like ECFP, EYFP, and etc. by cloning into the *Nhe*I and *Bsrg*I sites without the change of any amino acid of the linker between AcGFP1 and AGT.

The linker between AcGFP and the SNAP-tag (AGT) was modified by QuickChange (site-directed mutagenesis) with the oligonucleotides 5'-CTGTACAAGGCTCTAGCCGC**TGCA**GATATCAAGCTTACCATGGACAAAG-3' and 5'-CTTTGTCCATGGTAAGCTTGATATC**TGCA**GCGGCTAGAGCCTTGTACAG-3' to yield **A**. The resulting 10 amino acid linker was Ala-Leu-Ala-Ala-Ala-Asp-Ile-Lys-Leu-Thr.

CMV-MTS-Cer-AGT (B). Cerulean was amplified from pmscvcerulean (generous gift of doyle group) with the oligonucleotides 5'-CTAATTGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAG-3' and 5'-CTAGTGCTTGTACAGCTCGTCCATGCC-3' and inserted into the *Nhe*I and *Bsrg*I sites of pAcGFP1-C1 (clontech) yielding **pCer-C1. PCer-C1** was digested with *Age*I and *Bsrg*I and the resulting fragment inserted into the corresponding restriction sites of **A** to obtain **B**.

CMV-AcGFP1-AGT-NLS (E). The AcGFP-AGT fragment was amplified from **A** with oligonucleotides 5'-CTAATTGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGC-3'

and 5’-TGAACATCCGGAACCCAGCCCAGGCTTGC-3’ and ligated into the *Kpn*2I and *Nhe*I sites of pECFP-nuc (clontech) resulting in **E**.

CMV-AcGFP-TGN38-LS-1 (F). The 5’-phosphorylated oligonucleotides 5’-GATCTCGACGGCCGAAAGCGAGCGATTATCAGCGCCTGTGAA and 5’-AGCTTTCACAGGCGCTGATAATCGCTCGCTTTCGGCCGTCGA encoding RRPKASDYQRL-stop were annealed to each other and inserted into the *Bgl*II and *Hind*III sites of pAcGFP1-C1 (clontech) to give **F**.

CMV-AcGFP-TGN38-LS-2 (G). The 5’-phosphorylated oligonucleotides 5’-GATCTCGACGGCCGAAAGCGAGCGATTATCAGCGCCTGA and 5’-AGCTTCAGGCGCTGATAATCGCTCGCTTTCGGCCGTCGA encoding RRPKASDYQRL were annealed to each other and inserted into the *Bgl*II and *Hind*III sites of pAcGFP1-C1 (clontech) to give **G**.

CMV-AcGFP-TGN38-LS-3 (H). F was digested by *Bgl*II and blunted by the exonuclease Mung Bean Nuclease. The 5’-phosphorylated oligonucleotides 5’-TACCATCGCGACGACCGAAAGCGAGCGATTATCAGCGCCT and 5’-AGGCGCTGATAATCGCTCGCTTTCGGTCGTCGCGATGGTA encoding RRPKASDYQRL were annealed to each other and ligated into the blunted **F** to give **H**.

CMV-AcGFP-TGN38-LS-4 (I). F was digested by *Bgl*II and blunted by the exonuclease Mung Bean Nuclease. The 5’-phosphorylated oligonucleotides 5’-AGAGAAAGATCATAGCATTCGCTCTAGAAGGCAAACGCAGCAAAGTGAC and 5’-GTCACTTTGCTGCGTTTGCCTTCTAGAGCGAATGCTATGATCTTTCTCT encoding

KRKIIAFALEGKRSKVT were annealed to each other and ligated into the blunted **F** to give **I**.

CMV-AcGFP-TGN38-LS-5 (J). PAcGFP1-C1 was digested by *Bgl*II and blunted by the exonuclease Mung Bean Nuclease. The 5'-phosphorylated oligonucleotides 5'-GCGACTACCAGCGCCTGAGCGACTACCAGCGCCTGTAAA-3' and 5'-AGCTTTTACAGGCGCTGGTAGTCGCTCAGGCGCTGGTAGTCGC-3' were annealed to each other and ligated into the blunted pAcGFP1-C1 to give **J**.

CMV-Intron-Intein^{C}-AcGFP-TGN38 (K). An Intein^{C}-AcGFP-TGN38 fragment was amplified from **Q** with oligonucleotides 5'-ATATGACTCGAGCATGGGCACTAGTAGCACAGG-3' and 5'-TATACGTCTAGATTAAAGCTTTAGGTTCAAACGTTGG-3' and inserted into the *Xho*I and *Xba*I of pCI (Promega).

CMV-Intron-Intein^{C}-Cer-TGN38 (L). An Intein^{C}-Cer-TGN38 fragment was amplified from **R** with oligonucleotides 5'-ATATGACTCGAGCATGGGCACTAGTAGCACAGG-3' and 5'-TATACGTCTAGATTAAAGCTTTAGGTTCAAACGTTGG-3' and inserted into the *Xho*I and *Xba*I of pCI (Promega).

CMV-AcGFP-AGT-TGN38 (M). An AcGFP-AGT fragment was amplified from A with oligonucleotides 5'-CTAATTGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGC-3' and 5'-

TGAACAACCGGTACCCAGCCCAGGCTTGC-3' and cloned into the *Age*I site of **CMV-AcGFP-TGN38**.

T7-MTS-AcGFP-AGT (N). PET 11a was modified by PCR according to the Quickchange protocol with the oligonucleotides 5'-GGTGGCAGCAGCCAACTC**G**AGCTTCCTTTCGGGCTTTG-3' and 5'-CAAAGCCCGAAAGGAAGCT**C**GAGTTGGCTGCTGCCACC-3' to give **pET 11a'**. An MTS-AcGFP-AGT fragment was obtained by digestion of **A** with *Nhe*I and *Xho*I and inserted into the corresponding restriction sites of pET 28a.

T7-His-MTS-AcGFP-AGT (O). An MTS-AcGFP-AGT fragment was obtained by digestion of **A** with *Nhe*I and *Xho*I and inserted into the corresponding restriction sites of pET 28a.

T7-His-MTS-Cer-AGT (P). A cerulean fragment was obtained by digestion of **pCer-C1** with *Age*I and *Bsr*gI and inserted into the corresponding restriction sites of **O**.

T7-Intein^{C}-AcGFP-TGN38-His (Q). An Intein^{C} fragment was amplified from pTwin1 (NEB) with oligonucleotides 5'-ATACCATGGGCACTAGTAGCACAGGAAAAAGAGTTTC-3' and 5'-GGTGAAGCTAGCGCTGTTGTGTACAATGATGTCATTCGC-3' and cloned into the

*Nhe*I and *Nco*I sites of pET 28a affording **pET 28a-Intein^{C}**. A second *Bsrg*I restriction site, within the AcGFP1 gene of pAcGFP1-C1, was removed by silent mutation via site-directed mutagenesis according to the QuickQhange protocol (Stratagene) with the oligonucleotides 5'- CTACAACGCCCACAATGTGTA**T**ATCATGACCGACAAGGCC-3' and 5'- GGCCTTGTCGGTCATGAT**A**TACACATTGTGGGCGTTGTAG-3' to yield plasmid **pAcGFP1-C1'**. A TGN38 fragment was amplified from TGN38-pHluorin (generous gift from Terry E Machen; originally generated in the Miesenbock group) with oligonucleotides 5'-AGCTTATGTACAAAACCGGTGGAAGCG-3' and 5'-CTTATCATGTCTGCTCGAAGCG-3' and inserted into the BsrgI and SalI sites of **pAcGFP1-C1'** affording **CMV-AcGFP-TGN38**. An AcGFP-TGN38 fragment was amplified from **CMV-AcGFP-TGN38** with 5'-TTACTAACTAGTATGGTGAGCAAGGGCGC-3' and 5'-CCTCTACAAATGTGGTATGGCTG-3', digested with *Spe*I and *Hin*dIII and cloned into the *Nhe*I and *Hin*dIII sites of **pET 28a-Intein^{C}** in order to obtain **Q**.

T7-Intein^{C}-Cer-TGN38-His (R). Cerulean was amplified from pmscvcerulean (generous gift of doyle group) with the oligonucleotides 5'-CTAATTGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAG-3' and 5'-CTAGTGCTTGTACAGCTCGTCCATGCC-3' and inserted into the *Nhe*I and *Bsrg*I sites of **CMV-AcGFP-TGN38** yielding **CMV-Cer-TGN38**. A Cer-TGN38 fragment was amplified from **CMV-Cer-TGN38** with 5'-TTACTAACTAGTATGGTGAGCAAGGGCGC-3' and 5'-CCTCTACAAATGTGGTATGGCTG-3', digested with *Spe*I and *Hin*dIII and cloned into the *Nhe*I and *Hin*dIII sites of **pET 28a-Intein^{C}** in order to obtain **R**.

T7-Intein^{C}-AcGFP-His (S). An AcGFP fragment was amplified from **pAcGFP1-C1'** with the oligonucleotides 5'-TTACTAACTAGTATGGTGAGCAAGGGCGC-3' and 5'-CCTCTACAAATGTGGTATGGCTG-3', digested with *Spe*I and *Sal*I and inserted into the *Nhe*I and *Sal*I sites of **pET 28a-intein^{C}**.

T7-Intein^{C}-Cer-His (T). A cerulean fragment was amplified from **pCer-C1** with the oligonucleotides 5'-TTACTAACTAGTATGGTGAGCAAGGGCGC-3' and 5'-CCTCTACAAATGTGGTATGGCTG-3', digested with *Spe*I and *Sal*I and inserted into the *Nhe*I and *Sal*I sites of **pET 28a-intein^{C}**.

4.8.3 Synthesis

4-(Aminomethyl)-benzyl alcohol[21] **4-2**, 2,2,2-Trifluoro-*N*-(4-hydroxymethyl-benzyl)-acetamide[11] **4-3**, 1-(2-Amino-7H-purin-6-yl)-1-methyl-pyrrolidinium chloride[11] **4-5** and 3',6'-Dipyrrolidino-6-carboxyrhodamine[35,22] **4-8**, were synthesized following the published procedures.

***N*-[4-(2-Amino-*9H*-purin-6-yloxymethyl)-benzyl]-2,2,2-trifluoro-acetamide (4-6).[11]** NaH (95 %) (36 mg; 1.5 mmol) was added to a stirred solution of **4-3** (90.1 mg; 0.386 mmol) in 1 mL anhydrous DMF under argon atmosphere. Guanine derivative **4-5** (99.4 mg; 0.415 mmol) and DMAP (6.0 mg; 49 µmol) were added and the solution stirred at room temperature for 2 h. Water (0.1 mL) was added to quench the reaction and the solvent removed in vacuo. The crude product was purified by flash chromatography on

silica gel (CH_2Cl_2/MeOH 10:1; Rf = 0.26) providing **4-6** as white powder (100 mg, 68 %). ^{1}H NMR (DMSO-d_6, 400 MHz): δ 12.40 (s, 1H); 10.00 (s, 1H); 7.80 (s, 1H); 7.48 (d, J = 8.2 Hz, 2H); 7.29 (d, J = 8.2 Hz, 2H); 6.29 (s, 2H); 5.45 (s, 2H); 4.39 (d, J = 6.0 Hz, 2H).

***O*^{6}-(4-Aminomethyl-benzyl)guanine 4-7.**[11] Methylamine (3 mL; 33 % in EtOH) was added to a suspension of **4-6** in 1.5 mL methanol. The resulting solution was stirred at room temperature for 24 h. Concentration under reduced pressure yielded the product as white powder (93 mg, 100 %). ^{1}H NMR (DMSO-d_6, 400 MHz): δ 7.81 (s, 1H); 7.42 (d, J = 8.2 Hz, 2H); 7.33 (d, J = 8.2 Hz, 2H); 6.27 (s, 2H); 5.44 (s, 2H); 3.71 (s, 2H).

3',6'-Dipyrrolidino-6-amidobenzylguanine rhodamine (BGRH) 4-9.[36] HOBT (3.0 mg, 22 µmol) and DCC (13.0 mg, 63.0 µmol) were added to a solution of **4-8** (5.0 mg, 10 µmol) in 0.5 mL anhydrous DMF at 0 ℃ and stirred for 30 minutes. Benzyl guanine **4-7** (4.0 mg, 15 µmol) was added to the solution and stirred for 30 minutes. The solution was allowed to reach room temperature and stirred for 24 h. The organic solvent was removed under reduced pressure and the residue purified by flash chromatography on silica gel (CH_2Cl_2/MeOH 10:1) providing **4-9** as red powder (5.2 mg, 68 %).

Rhodamine B base[23] **4-11**, rhodamine B piperazine amide[23] **4-12** and rhodamine B 4-(3-carboxypropionyl)piperazine amide[23] **4-13** and glycine *S*-benzyl ester, hydrochloride salt[37] **4-14** were synthesized following the published procedure.

Rhodamine B 4-(3-amidopropionyl, 4-glycine S-benzyl ester)piperazine amide 4-15. HOBT (25.0 mg, 0.185 mmol) and EDCI (31.2 mg, 0.163 mmol) were added to a solution of **4-13** (100 mg, 0.163 mmol) in 2 mL DMF at 0 ℃ and stirr ed for 30 minutes. $NaHCO_3$ (12.0 mg, 0.141 mmol) and **4-14** (35.6 mg, 0.163 mmol) were added and the solution stirred for 1 h. The solution was allowed to reach room temperature and stirred for 24 h. The organic solvent was removed under reduced pressure and the residue purified by flash chromatography on silica gel (CH_2Cl_2/MeOH 10:1) providing **4-15** as red powder (85 mg, 67 %). ^{1}H NMR (CD_3OD, 400 MHz): δ 7.79-7.76 (m, 2H); 7.71-7.69 (m, 1H); 7.53-7.50 (m, 1H); 7.28 (d, *J* = 9.6 Hz, 2H); 7.25-7.19 (m, 5H); 7.08-7.05 (m, 2H); 6.95 (d, *J* = 2.2 Hz; 2H); 4.07 (s, 2H); 4.05 (s, 2H); 3.67 (q, *J* = 7.2 Hz, 8H); 3.41 (br s, 8H); 2.60 (br s, 2H); 2.55 (br s, 2H), 1.30 (t, *J* =7.2 Hz, 12H). ESI-MS: *m/z* 774.4 [M^+]. ESI-HRMS calculated for [M^+] C_{45} H_{52} N_5 O_5 S_1 774.3689, found 774.3694.

Rhodamine B split-intein^N 4-16.[38] A 20 mM solution of Rhodamine derivative **4-15** (161 µL, 3.23 µmol) in DMF was added to a solution of the polypeptide CISGDSLISLASR (3.23 µmol) in 2.15 mL aqueous buffered solution (6 M guanidinium hydrochloride, 100 mM Na-phosphate pH 7.60, 4 % thiophenol) and stirred at room temperature for 24 h. The crude product was purified by reversed phase HPLC (5.1 mg, 80 %). MALDI-MS: *m/z* 1971.0 [M^+]. MALDI-HRMS calculated for [M^+] C_{92} H_{140} N_{21} O_{25} S_1 1971.0044, found 1971.0170.

Benzyl guanine zinquin 4-18. HOBT (11.0 mg, 81.4 µmol) and EDCI (16.0 mg, 83.5 µmol) were added to a solution of **4-17** (27.3 mg, 70.6 µmol) in 2 mL anhydrous THF at

0 ℃ and stirred for 30 minutes. Benzyl guanine **4-7** (18.8 mg, 69.6 µmol) was dissolved in 2 mL anhydrous DMF and added dropwise to the solution and stirred for 30 minutes. The solution was allowed to reach room temperature and stirred for 24 h. The organic solvent was removed under reduced pressure and the residue purified by flash chromatography on silica gel (CH_2Cl_2/MeOH 10:1) providing **4-17** as off-white powder (22 mg, 50 %). FAB-HRMS calculated for [$M+H^+$] C_{32} H_{31} N_8 O_5 S_1 639.21392, found 639.21381.

N-{Tris{3-{2-(2-benzyloxy-phenyl)-methyl-4,5-dihydrooxazolyl]propyloxymethyl}-methyl-4-carboxypropionamide **4-20** was synthesized following the published procedures.[26]

3-methyl isonitrosoacetanilide[29] **4-22**, 6-methyl isatin[30] **4-23**, 4-methyl isatin[30] **4-24**, 4-methyl isonitrosoacetanilide[29] **4-26**, 5-methyl isatin[30] **4-27** and 6,6'-dimethyl-2,2'-bicinchoninic acid **4-28** were synthesized following the published procedures.

6,6'-Dimethyl-2,2'-bicinchoninic acid 4-28.[34] A mixture of **4-27** (4.02 g, 24.8 mmol), acetoin (1.15 g, 13.0 mmol) and 20 g of 33 % KOH-solution (w/w) was heated at 100 ℃ for 24 h. The suspension was allowed to cool, the K-salt filtered and dissolved in hot water. The product was precipitated by acidification with acetic acid, filtered and dried obtaining **4-28** as white powder (1.74 g, 38 %). ^{1}H NMR (DMSO-d_6, 400 MHz): δ 8.69 (s, 2H); 8.49 (s, 2H); 7.98 (d, J = 8.6 Hz, 2H); 7.54 (d, J = 8.6 Hz, 2H); 3.32 (s, 6H). EI-HRMS calculated for [M^+] C_{22} H_{16} N_2 O_4 372.11101, found 372.10893.

6,6'-Dimethyl-2,2'-bicinchoninic acid dichloride 4-29. Bicinchoninic acid derivative **4-28** (250 mg, 0.672 mmol) was dissolved in 5 mL $SOCl_2$ and heated under reflux overnight. Thionyl chloride was removed under reduced pressure to yield **4-29** (325 mg, 100 %). The product was used without any further purification.

6,6'-Dimethyl-2,2'-bicinchoninic acid diethylester 4-30. To a solution of NaOEt (3.04 mmol) in 10 mL ethanol **4-29** (250 mg, 0.611 mmol) was added and the solution stirred at room temperature for 24 h. After concentration of the solution water (20 mL) was added and the solution extracted three times with 30 mL $CHCl_2$. The organic layers were combined and the solvent removed under reduced pressure to yield **4-30** as white powder (130 mg, 50 %). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 9.24 (s, 2H); 8.56 (s, 2H); 8.21 (d, *J* = 8.6 Hz, 2H); 7.65 (d, *J* = 8.6 Hz, 2H); 4.59 (q, *J* = 7.2 Hz, 4H); 2.62 (s, 6H); 1.55 (t, J = 7.2 Hz, 6H). EI-HRMS calculated for [M^+] C_{26} H_{24} N_2 O_4 428.17361, found 428.17078.

6-Methyl-6'-bromomethyl-2,2'-bicinchoninic acid diethylester 4-31. To a 10 mL CCl_4 solution of **4-30** (113 mg, 263 µmol) was added NBS (46.7 mg, 263 µmol) and a catalytic amount of benzoyl peroxide (≈ 3 mg). The mixture was heated under reflux for 24 h and concentrated under reduced pressure. The crude product was purified by chromatography on silica gel (Hexanes/EtOAc 20:1) providing **4-31** as white powder (39.9 mg, 30 %). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 9.31 (s, 1H); 9.23 (s, 1H); 8.85 (s, 1H); 8.56 (s, 1H); 8.30 (d, *J* = 8.6 Hz, 1H); (8.22 (d, *J* = 8.6 Hz, 1H); 7.85 (d, *J* = 8.6 Hz, 1H);

7.66 (d, *J* = 8.6 Hz, 1H); 4.73 (s, 2H); 4.59 (q, *J* = 7.2 Hz, 4H); 2.62 (s, 3H); 1.55 (t, J = 7.2 Hz, 6H). EI-HRMS calculated for [M^+] C_{26} H_{23} Br N_2 O_4 506.08412, found 506.07996.

6,6'-Dimethyl-2,2'-bicinchoninic acid di-*tert*-butylester 4-34. EDCI (1.567 g, 8.12 mmol) was added to a solution of **4-28** (1.375 g, 3.69 mmol), DMAP (0.451 g, 3.69 mmol) and *t*-BuOH (0.629 g, 8.49 mmol) in 50 mL CH_2Cl_2 at 0 ℃ and stirred for 2 h. The solution was allowed to reach room temperature and stirred for 24 h. The organic solvent was removed under reduced pressure and the residue purified by flash chromatography on silica gel (Hexanes/EtOAc 7:3) providing **4-34** as white powder (492 mg, 28 %). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 9.14 (s, 2H); 8.51 (s, 2H); 8.19 (d, *J* = 8.6 Hz, 2H); 7.53 (d, *J* = 8.6 Hz, 2H); 2.61 (s, 6H); 1.75 (s, 18H). EI-HRMS calculated for [M^+] C_{30} H_{32} N_2 O_4 484.23621, found 484.23364.

6-Methyl-6-bromomethyl-2,2'-bicinchoninic acid di-*tert*-butylester 4-35. To a 30 mL CCl_4 solution of **4-34** (480 mg, 990 µmol) was added NBS (176 mg, 991 µmol) and a catalytic amount of benzoyl peroxide (≈ 10 mg). The mixture was heated under reflux for 24 h and concentrated under reduced pressure. The crude product was purified by chromatography on silica gel (Hexanes/EtOAc 30:1) providing **4-35** as white powder (124 mg, 28 %). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 9.21 (s, 1H); 9.13 (s, 1H); 8.81 (s, 1H); 8.51 (s, 1H); 8.28 (d, *J* = 8.6 Hz, 1H); 8.19 (d, *J* = 8.6 Hz, 1H); 7.82 (d, *J* = 8.6 Hz, 1H); 7.53 (d, *J* = 8.6 Hz, 1H); 4.73 (s, 2H); 2.62 (s, 3H); 1.76 (s, 18H). EI-HRMS calculated for [M^+] C_{30} H_{31} Br N_2 O_4 562.16472, found 562.14395.

2-Bromomethyl-9-methyl-4,7-diphenyl-1,10-phenanthroline 4-38. To a 10 mL CCl_4 solution of **4-37** (100 mg, 277 µmol) was added NBS (49.4 mg, 277 µmol) and a catalytic amount of benzoyl peroxide (≈ 3 mg). The mixture was heated under reflux for 24 h and concentrated under reduced pressure. The crude product was purified by flash chromatography on silica gel (Hexanes/EtOAc 7:3; R_f = 0.24) providing **4-38** as white powder (49.3 mg, 41 %). ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.85-7.78 (m, 4H); 7.55-7.49 (m, 10H); 5.03 (s, 2H); 3.01 (s, 3H). ^{13}C NMR ($CDCl_3$, 400 MHz): δ 159.0; 156.4; 149.6; 148.6; 145.9; 145.3; 137.8; 137.7; 129.58; 129.53; 128.53; 128.51; 128.48; 128.4; 125.8; 124.8; 124.4; 124.2; 123.6; 122.7; 35.1; 25.92. EI-HRMS calculated for [M^+] C_{26} H_{19} Br N_2 438.07316, found 438.07348.

4.8.4 Analysis of FRET Pair MTS-AcGFP-AGT-Rhodamine

The properties of the FRET pair were calculated utilizing equations 1-5.[39] The molar extinction coefficient of the acceptor rhodamine was derived from the absorption spectrum by the Lambert-Beer equation, in which *A* is the absorption intensity, *c* the concentration and *l* the path length (equation 1). The overlap integral *J* (*λ*) was determined via the integrals of the fluorescence intensities of the donor F_D (equation 2). The Foerster radius R_0 was calculated with the parameters κ^2 (orientation of the transition dipoles in space = 2/3 for the assumed dynamic random averaging of donor and acceptor), *n* (refractive index = 1.4), Q_D (quantum yield) and overlap integral *J* (*λ*) (in units of M^{-1} cm^{-1} $(nm)^4$) (equation 3). The quantum yield of AcGFP (MTS-AcGFP-AGT) was determined by comparison of the integrated intensity of the unknown AcGFP with a

fluorescein standard solution in 0.1 M NaOH (resulting quantum yield of AcGFP Q_D = 57.3 %).[40,41] Equation 4 was used to calculate the energy transfer efficiency (FRET efficiency) E from the fluorescence intensity of the donor in the presence (F_{DA}) and absence (F_D) of energy transfer. The distance between donor and acceptor r was derived with the previously calculated parameters E and R_0 by equation 5.

$$\varepsilon(\lambda) = \frac{A}{c\,l} \qquad \text{(equation 1)}$$

$$J(\lambda) = \frac{\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda}{\int_0^\infty F_D(\lambda)d\lambda} \qquad \text{(equation 2)}$$

$$R_0 = 0.211(\kappa^2 n^{-4} Q_D J(\lambda))^{\frac{1}{6}} \qquad \text{(equation 3)}$$

$$E = 1 - \frac{F_{DA}}{F_D} \qquad \text{(equation 4)}$$

$$E = \frac{R_0^6}{R_0^6 + r^6} \qquad \text{(equation 5)}$$

Purification of expressed proteins: The proteins were purified either by a Ni^{2+}-NTA superflow column or by Ni^{2+}-NTA superflow batch purification according to the Quiagen protocol (The Quiagen expressionist volume 6).

UV-vis spectroscopy: UV-vis absorption spectra were recorded at 25 ℃ u sing a Varian Bio50 UV-vis spectrometer with constant-temperature accessory.

Fluorescence spectroscopy: Emission and excitation spectra were recorded with a PTI fluorimeter utilizing FELIX software. All fluorescence spectra have been corrected for the spectral response of the detection system (emission correction parameters provided by the manufacturer) and for the spectral irradiance of the excitation channel by a calibrated photodiode.

NMR spectroscopy: All NMR spectra (^{1}H, ^{13}C) were recorded on either a Bruker DRX 500 or Mercury VX 400.

4.9 Literature

(1) Richmond, T. A.; Takahashi, T. T.; Shimkhada, R.; Bernsdorf, J. *Biochemical and Biophysical Research Communications* **2000**, *268*, 462-465.

(2) Petrat, F.; Weisheit, D.; Lensen, M.; de Groot, H.; Sustmann, R.; Rauen, U. *Biochemical Journal* **2002**, *362*, 137-147.

(3) Sandell, E. B. *Colorimetric determination of traces of metals*; Interscience Publishers, inc.: New York, N.Y.,, 1944.

(4) Stokes, H. N.; Cain, J. R.; United States. Bureau of standards. *On the colorimetric determination of iron with special reference to chemical reagents*; Gov't print. off.: Washington,, 1907.

(5) Herd, S. M.; Camakaris, J.; Christofferson, R.; Wookey, P.; Danks, D. M. *Biochemical Journal* **1987**, *247*, 341-347.

(6) Williams, R. J. P.; da Silva, J. *Coordination Chemistry Reviews* **2000**, *200*, 247-348.

(7) Breuer, W.; Shvartsman, M.; Cabantchik, Z. I. *International Journal of Biochemistry & Cell Biology* **2008**, *40*, 350-354.

(8) Ludwig, C.; Pfeiff, M.; Linne, U.; Mootz, H. D. *Angewandte Chemie-International Edition* **2006**, *45*, 5218-5221.

(9) Sun, W. C.; Yang, J.; Liu, X. Q. *Journal of Biological Chemistry* **2004**, *279*, 35281-35286.

(10) Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 9955-9959.

(11) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nature Biotechnology* **2003**, *21*, 86-89.

(12) Valeur, B. *Molecular fluorescence : principles and applications*; Wiley-VCH: Weinheim ; New York, 2002.

(13) Lackie, J. M.; Dow, J. A. T.; Blackshaw, S. E. *The dictionary of cell biology*; 2nd ed.; Academic Press: London ; San Diego, 1995.

(14) Hicks, G. R.; Raikhel, N. V. *Annual Review of Cell and Developmental Biology* **1995**, *11*, 155-188.

(15) Miesenbock, G.; De Angelis, D. A.; Rothman, J. E. *Nature* **1998**, *394*, 192-195.

(16) Humphrey, J. S.; Peters, P. J.; Yuan, L. C.; Bonifacino, J. S. *Journal of Cell Biology* **1993**, *120*, 1123-1135.

(17) Luzio, J. P.; Brake, B.; Banting, G.; Howell, K. E.; Braghetta, P.; Stanley, K. K. *Biochemical Journal* **1990**, *270*, 97-102.

(18) The transfections and microscopy was carried out by Reagan McRae

(19) Pap, E. H. W.; Dansen, T. B.; van Summeren, R.; Wirtz, K. W. A. *Experimental Cell Research* **2001**, *265*, 288-293.

(20) Buchman, A. R.; Berg, P. *Molecular and Cellular Biology* **1988**, *8*, 4395-4405.

(21) Lee, J.; Kang, M.; Shin, M.; Kim, J. M.; Kang, S. U.; Lim, J. O.; Choi, H. K.; Suh, Y. G.; Park, H. G.; Oh, U.; Kim, H. D.; Park, Y. H.; Ha, H. J.; Kim, Y. H.; Toth, A.; Wang, Y.; Tran, R.; Pearce, L. V.; Lundberg, D. J.; Blumberg, P. M. *Journal of Medicinal Chemistry* **2003**, *46*, 3116-3126.

(22) Woodroofe, C. C.; Lim, M. H.; Bu, W. M.; Lippard, S. J. *Tetrahedron* **2005**, *61*, 3097-3105.

(23) Nguyen, T.; Francis, M. B. *Organic Letters* **2003**, *5*, 3245-3248.

(24) Sumalekshmy, S.; Henary, M. M.; Siegel, N.; Lawson, P. V.; Wu, Y.; Schmidt, K.; Bredas, J. L.; Perry, J. W.; Fahrni, C. J. *Journal of the American Chemical Society* **2007**, *129*, 11888-+.

(25) The synthesis was carried out by Dr. S. Sumalekshmy

(26) Kikkeri, R.; Traboulsi, H.; Humbert, N.; Gumienna-Kontecka, E.; Arad-Yellin, R.; Melman, G.; Elhabiri, M.; Albrecht-Gary, A. M.; Shanzer, A. *Inorganic Chemistry* **2007**, *46*, 2485-2497.

(27) The synthesis was carried out by Dr. Jing Li

(28) The calculations were carried out by Dr. Christoph Fahrni.

(29) Su, W. K.; Zhong, W. H.; Bian, G. F.; Shi, X. J.; Zhang, J. P. *Organic Preparations and Procedures International* **2004**, *36*, 499-547.

(30) Marvel, C. S.; Hiers, G. S. *Organic Syntheses* **1925**, *5*, 71-74.

(31) Grimshaw, J.; Begley, W. J. *Synthesis-Stuttgart* **1974**, 496-498.

(32) Sadler, P. W. *Journal of Organic Chemistry* **1956**, *21*, 169-170.

(33) Sandmeyer, T. *Helvetica Chimica Acta* **1919**, *2*, 234-242.

(34) Lesesne, S. D.; Henze, H. R. *Journal of the American Chemical Society* **1942**, *64*, 1897-1900.

(35) The synthesis was carried out by Dr. Jing Li

(36) The synthesis was carried out by Dr. Jing Li

(37) Alsina, J.; Yokum, T. S.; Albericio, F.; Barany, G. *Journal of Organic Chemistry* **1999**, *64*, 8761-8769.

(38) Schuler, B.; Pannell, L. K. *Bioconjugate Chemistry* **2002**, *13*, 1039-1043.

(39) Lakowicz, J. R. *Principles of fluorescence spectroscopy*; 3rd ed.; Springer: New York, 2006.

(40) Demas, J. N.; Crosby, G. A. *Journal of Physical Chemistry* **1971**, *75*, 991-&.

(41) Magde, D.; Wong, R.; Seybold, P. G. *Photochemistry and Photobiology* **2002**, *75*, 327-334.

CHAPTER V

OUTLOOK

The enantioselective catalysis of the cyclopropanation of styrene with EDA by semi-synthetic proteins, based on myoglobin reconstituted with ruthenium porphyrins, failed (see chapter II and chapter III). The catalytic activity of ruthenium porphyrin ceased upon reconstitution of the apoprotein. Therefore the deactivation of the catalyst has to be due to the interaction of the synthetic cofactor with the residues of the active site. The successful reconstitution of myoglobin with a synthetic cofactor was confirmed by UV-Vis spectroscopy; however, no structural information, such as orientation of the cofactor in the active site pocket, can be gained from spectral shifts in UV-Vis spectra. The orientation of the synthetic cofactor that is different from the orientation of native heme cofactor may have led to the inhibition of the cyclopropanation reaction. Therefore the structure of the cofactor embedded in the active site pocket data can be obtained by X-ray crystallography and reveal any differences to the native heme pocket. Differences from the native myoglobin active site are expected to be more pronounced in case of the myoglobin single mutants (H64, H93) and especially the double mutant (H64G/H93G) due to the removal of hydrogen bond donors and the increase of the size of the active site pocket. Besides, flaws in the architecture of the active site, such as a partial collapse of the active site pocket caused by the mutations, can be detected by the X-ray structures. Hence, X-ray structures can direct future efforts to convert myoglobin,

reconstituted with ruthenium porphyrins, to an active cyclopropanase. Activity may be induced by site-directed mutagenesis based on the X-ray crystallography data.

A 1:1 stochiometry (ligand:metal) of the metal-chelator domain is required for the successful sensing of labile metal pools of copper, iron, and zinc by semi-synthetic proteins (see chapter IV). All attempts to failed to generate a chelator that forms Cu(I)-complexes with the necessary 1:1 stochiometry. The design of the desired copper chelator was based on the linking of bathocuproine or bicinchoninic acid derivatives with a diethanolamine linker by the formation of an alkoxide and subsequent reaction with a halide. Other methods to connect these copper(I) chelators may be more successful. Alternative reaction designs should be based on established coupling reactions (usually palladium based reactions), such as the Heck or Sonogashira reaction.

A ratiometric zinc(II)-chelator (with a 1:1 stochiometry) was attached to an AGT based protein and the resulting semi-synthetic protein exhibited strong changes in fluorescence in the presence of zinc(II). These *in vitro* studies represent an imperative step towards the goal of *in vivo* cell imaging of labile zinc(II) pools. The protein-chelator conjugate will be used to study labile zinc(II) pools in mitochondria, the endoplasmatic reticulum, the Golgi apparatus, and the nucleus. Consequently, mammalian cells have to be transfected with plasmids encoding the AGT with the appropriate localization sequences. These cells are subsequently incubated with the zinc(II) chelator upon which the semi-synthetic protein is formed *in vivo* and the conjugate directed to the subcellular location. The semi-synthetic protein can also be generated *in vitro* by over-expressing the protein domain in bacteria and reacting the purified protein with the zinc(II) chelator.

The formed fluorescent chelator-protein conjugate can be introduced into the cell via microinjection.